

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:14:42 : Search time 87 Seconds
(without alignments)
490.775 Million cell updates/sec

Title: US-09-852-100A-2

Perfect score: 1439

Sequence: 1 MHILKSPNVPRAHGQKNT.....TLRLSLITNEIRKLTOLYP 269

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

```
1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1439	100.0	269	AAW94291	Human beta-amyloid
2	1439	100.0	269	AAV70759	Human beta-amyloid
3	1439	100.0	269	AAE33877	Human BBP-1 protei
4	748	52.0	139	AAV12358	Human 5' EST secre
5	673.5	46.8	162	AAV36021	Extended human sec
6	667.5	46.4	148	AAV12426	Human 5' EST secre
7	338	23.5	178	ABB5236	Drosophila melanog
8	293	20.4	100	AAU97631	rRNA polymerase II
9	201	14.0	221	ABR47818	Human secreted pro

10	201	14.0	221	24	ABR00112	Human gene 102 enc
11	201	14.0	222	20	AAV27658	Human secreted pro
12	200	13.9	221	21	AAV70761	Human beta-amyloid
13	200	13.9	221	24	ABR48493	Human Amyloid Apop
14	200	13.9	247	21	AAE08641	A human TANGO 223
15	200	13.9	247	21	AAE08656	A human TANGO 223
16	200	13.9	247	21	AAE08657	A human TANGO 223
17	200	13.9	247	21	AAE08658	A human TANGO 223
18	196	13.6	230	21	AAE08645	DNA encoding a mur
19	182	12.6	284	22	ABE59014	Drosophila melanog
20	152.5	10.6	225	22	ABB11574	Human secreted pro
21	138.5	9.6	214	21	AAV87351	Human signal pepti
22	136.5	9.5	214	21	AAV70760	Human beta-amyloid
23	136.5	9.5	224	21	AAE43573	Human cancer assoc
24	132	9.2	224	22	ABE65466	Drosophila melanog
25	121	8.4	171	22	AAE67818	Human guanosine tri
26	121	8.4	171	22	AAE93475	Human polypeptide,
27	121	8.4	171	22	AAE93578	Human polypeptide,
28	121	8.4	171	22	AAE81266	Human AFP protei
29	121	8.4	171	23	ABE90376	Human polypeptide
30	121	8.4	172	20	AAE45270	Human secreted pro
31	119	8.3	140	23	ABE40585	Human secreted pro
32	115.5	8.0	213	23	ABE72307	Rat protein isolat
33	95.5	6.6	944	24	AAE32044	Human vesicle-assc
34	90.5	6.3	341	22	AAE41757	Human polypeptide
35	88.5	6.2	224	22	ABE20306	Novel human diagn
36	88	6.1	68	24	AAE33878	Human BBP-1 protei
37	88	6.1	764	20	AAE94360	Human Factor B and
38	86.5	6.0	316	22	ABE64318	Drosophila melanoc
39	86	6.0	764	20	AAE94365	Human Factor B and
40	86	6.0	764	20	AAE94372	Human Factor B and
41	85	5.9	310	17	AAE05156	Rat mature LAMP.
42	85	5.9	338	17	AAE05153	Rat LAMP residues
43	85	5.9	338	17	AAE05154	Rat LAMP residues
44	85	5.9	576	21	AAE75498	Neisseria gonorrhoe
45	85	5.9	576	24	ABE80808	N. gonorrhoeae am

ALIGNMENTS

```
RESULT 1
AAW94291
ID AAW94291 standard; Protein; 269 AA.
XX
XX AAW94291;
AC AAW94291;
XX
XX 27-APR-1999 (first entry)
DE Human beta-amyloid peptide-binding protein (BBP).
XX
XX Beta-amyloid peptide binding protein; BBP; beta-amyloid protein; BAP;
XX human; Alzheimer's disease.
XX
XX Homo sapiens.
CS
XX
XX Key Location/Qualifiers
XX Region 68..269
XX FT /note- "specifically claimed fragment having
XX FT beta-amyloid peptide binding activity."
XX
XX W05846636-A2.
XX
XX PD 22-OCT-1998.
XX
XX PF 14-APR-1998; 98WC-US07462. ✓
XX
XX PR 16-APR-1997; 97US-0064583.
XX
XX (AMHP ) AMERICAN HOME PROD CORP.
XX
XX PA Bard JA, Jacobsen JS, Kajkowski EM, Ozenberger BA;
XX Walker SG;
```

```

XX WPI: 1999-080736/G7.
DR N-PSDB: AAX05735.
XX
XX Polynucleotide encoding beta-amyloid peptide binding protein - used
PT to identify inhibitors of beta-amyloid peptide for treating
PT Alzheimer's disease
XX
XX Claim 7: Pages 43-44; 59pp: English.
XX
CC The present sequence represents a beta-amyloid peptide binding protein
CC (BBP). The polynucleotide comprising the entire BBP nucleotide sequence
CC of clone BBP1-f1 is deposited under the accession number ATCC 98617. The
CC polynucleotide comprising a fragment of BBP (nucleotides 202-807 of the
CC full length BBP) of clone PEK196 is deposited as ATCC 98395. Host cells
CC transformed with a vector comprising the BBP nucleic acid are used for
CC the recombinant production of the protein. The protein can be used in a
CC method for diagnosing a disease characterised by aberrant expression of
CC human beta-amyloid protein (BAP). The protein can also be used in a
CC method for screening for compounds which regulate expression of a BAP
CC binding protein. The proteins, antibodies and identified compounds can be
CC used in the treatment or prevention of Alzheimer's disease.
XX
XX Sequence 269 AA:
XX
XX Query Match 100.0%; Score 1439; DB 20; Length 269;
XX Best Local Similarity 100.0%; Pred. No. 9,4e-142;
XX Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MHILKGSPNVIPIRAHGOKNTRDGTGLYPMRGPFKNLALLPFLGSGSGSGEKVSV 60
Db 1 MHILKGSPNVIPIRAHGOKNTRDGTGLYPMRGPFKNLALLPFLGSGSGSGEKVSV 60
QY 61 SKMAAAMPSPGSAPEAVTARLVGLWLFVSVTTGPGAVATSGAGEESLKCEDLKVGQYIC 120
Db 61 SKMAAAMPSPGSAPEAVTARLVGLWLFVSVTTGPGAVATSGAGEESLKCEDLKVGQYIC 120
QY 121 KDPKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFPPKPSICRNNG 180
Db 121 KDPKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFPPKPSICRNNG 180
QY 181 YSYKVAVALSLFLGWLADRFYLGYPALGLLKCTVGFCGIGSLIDFLILSMQIVGPSDG 240
Db 181 YSYKVAVALSLFLGWLADRFYLGYPALGLLKCTVGFCGIGSLIDFLILSMQIVGPSDG 240
QY 241 SSVIIDYGTGTRLTSLTNETFRKTQLYP 269
Db 241 SSVIIDYGTGTRLTSLTNETFRKTQLYP 269
XX
XX Location/Qualifiers
XX 177..198
XX /label= Transmembrane_domain_1
XX Domain
XX 199..201
XX /label= DRF_motif
XX /note= "Substitution of the Arg abrogates protection"
XX

```

```

FT Domain 213..238
FT /label= Transmembrane_domain_2
XX
XX WQ20002125-A2.
XX
XX 20-APR-2000.
XX
XX 13-OCT-1999; 99WG-US21521.
XX
XX 13-OCT-1998; 98US-0104104.
XX
XX (AMHP ) AMERICAN HOME PROD CORP.
XX
XX ozenberger BA, kajkowski EM, LO CF;
XX
XX WPI: 2000-317982/27.
XX N-PSDB: AAX52369.
XX
XX Novel G-protein-coupled receptor-like proteins and polynucleotides
XX useful for regulating apoptosis, comprises integral membrane protein
XX traversing the membrane twice.
XX
XX Example 1: Page 62-63; 68pp: English.
XX
XX The present sequence is the beta-amyloid peptide (BAP) binding protein-1
XX (BBP1). It is an integral membrane protein, that traverse the membrane
XX twice. It is related to G protein-coupled receptor (GPCR) protein
XX superfamily. It interacts with G-alpha proteins and regulates the
XX activity of G-protein signalling pathways. BBP genes are widely expressed
XX in neuronal cells of nonhuman primate (NHP) brain and overexpressed in
XX some tumours. It functions as a suppressor of apoptosis induction. BBP
XX proteins are used as immunogens to raise antibodies, useful as
XX therapeutics and as antigens in solid phase assays. They are also useful
XX as reagents to identify molecules which effect the interaction of BBP and
XX a cloned protein, that are useful in the treatment or prevention of
XX diseases associated with apoptosis. The polynucleotides are useful for
XX diagnostics.
XX
XX Note: In claim 5, the patent claims an amino acid sequence from figure 2.
XX However, figure 2 does not contain any sequence. It is inferred from the
XX disclosure that the figure 2 sequence refers to BBP1 protein, shown in
XX this sequence.
XX
XX Sequence 269 AA:
XX
XX Query Match 100.0%; Score 1439; DB 21; Length 269;
XX Best Local Similarity 100.0%; Pred. No. 9,4e-142;
XX Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MHILKGSPNVIPIRAHGOKNTRDGTGLYPMRGPFKNLALLPFLGSGSGSGEKVSV 60
Db 1 MHILKGSPNVIPIRAHGOKNTRDGTGLYPMRGPFKNLALLPFLGSGSGSGEKVSV 60
QY 61 SKMAAAMPSPGSAPEAVTARLVGLWLFVSVTTGPGAVATSGAGEESLKCEDLKVGQYIC 120
Db 61 SKMAAAMPSPGSAPEAVTARLVGLWLFVSVTTGPGAVATSGAGEESLKCEDLKVGQYIC 120
QY 121 KDPKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFPPKPSICRNNG 180
Db 121 KDPKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFPPKPSICRNNG 180
QY 181 YSYKVAVALSLFLGWLADRFYLGYPALGLLKCTVGFCGIGSLIDFLILSMQIVGPSDG 240
Db 181 YSYKVAVALSLFLGWLADRFYLGYPALGLLKCTVGFCGIGSLIDFLILSMQIVGPSDG 240
QY 241 SSVIIDYGTGTRLTSLTNETFRKTQLYP 269
Db 241 SSVIIDYGTGTRLTSLTNETFRKTQLYP 269
XX
XX RESULT 3
XX ID AAE33877
XX ID AAE33877 standard: Protein; 269 AA.
XX

```

```

AC AAE33877;
XX
XX
DT 02-MAY-2003 (first entry)
XX
XX DE Human BBP-1 protein.
XX
XX Human; beta-amyloid peptide-binding protein; BAP; Abeta; betaA2P; BBP;
KW Alzheimer's disease; AD; transgenic; transgenic animal; gene therapy;
KW neuroprotective; neotropic.
XX
XX Homo sapiens.
OS
XX WO200290499-A2.
XX
XX 14-NOV-2002.
XX
XX 06-MAY-2002; 2002WO-US14223.
XX
XX 09-MAY-2001; 2001US-0852100.
XX
XX (AMBP ) WYETH.
XX
XX Ozenberger BA, Bard JA, Kajkowski EM, Jacobsen JS, Walker SG;
PI Sofia HJ, Howland DS;
XX
XX WPI; 2003-120537/11.
XX
XX N-PSDB; AAD51940.
XX
XX New human beta-amyloid peptide-binding protein, useful for diagnosing
PT and/or treating diseases associated with aberrant expression of
PT beta-amyloid peptide, e.g. Alzheimer's disease.
XX
XX Claim 4; Page 84-85; 85pp; English.
XX
XX The present invention relates to novel human beta-amyloid peptide (BAP;
CC Abeta, betaA2P)-binding (BBP) proteins and polynucleotides encoding such
CC proteins. BBP sequences are useful to diagnose and/or treat diseases
CC associated with aberrant expression of human BAP such as Alzheimer's
CC disease (AD). They are used to generate transgenic animals. Sequences
CC of the invention are also used in gene therapy. The present sequence
XX is human BAP-1 protein.
XX
XX Sequence 269 AA:
XX
XX Query Match 100.0%; Score 1439; DB 24; Length 269;
XX Best Local Similarity 100.0%; Pred. No. 9.4e-142;
XX Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MHILKSPNVIIPRAGQKNIRRDGTCLYPMRGPPKKNLALLPFLPLGGSGSGSEKYSV 60
DQ 1 MHILKSPNVIIPRAGQKNIRRDGTCLYPMRGPPKKNLALLPFLPLGGSGSGSEKYSV 60
QY 61 SKMAAASFSGSPAPEAVTARLVGVLMFVSVTITGPGAVATAGGSESLKCEDLKVGVIC 120
DQ 61 SKMAAASFSGSPAPEAVTARLVGVLMFVSVTITGPGAVATAGGSESLKCEDLKVGVIC 120
QY 121 KDPKINDATQEPVNCNTNTAHVSCFPAPNITCKDSSGNETHFTGNEVGFFKPISCRNVNG 180
DQ 121 KDPKINDATQEPVNCNTNTAHVSCFPAPNITCKDSSGNETHFTGNEVGFFKPISCRNVNG 180
QY 181 YSYKVAVALSLFLGLGADRFYLVGYPALGLLKFCTVPGFCGSLIDFIIISMQIVGPSDG 240
DQ 181 YSYKVAVALSLFLGLGADRFYLVGYPALGLLKFCTVPGFCGSLIDFIIISMQIVGPSDG 240
QY 241 SSIYIDYVGTGTRLTSLISINETFRKTLQYP 269
DQ 241 SSIYIDYVGTGTRLTSLISINETFRKTLQYP 269

```

```

RESULT 4
AA12358
ID AAY12358 standard; Protein: 139 AA.
XX

```

```

AC AAY12358;
XX
XX 17-JUN-1999 (first entry)
XX
XX Human 5' EST secreted protein SEQ ID NO:389.
XX
XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.
XX
XX Homo sapiens.
XX
XX WO9906548-A2.
XX
XX 11-FEB-1999.
XX
XX 31-JUL-1998; 98WO-IB01222.
XX
XX 01-AUG-1997; 97US-0905135.
XX
XX (GEST ) GENSET.
XX
XX Ducicert A, Dumas Milne Edwards J, Lacroix R;
XX
XX WPI; 1999-153778/13.
XX
XX N-PSDB; AAX41191.
XX
XX New nucleic acids encoding human secreted proteins - obtained from
PT cDNA libraries prepared from e.g. liver, ovary, brain, prostate,
PT kidney, lung, umbilical cord, placenta and colon tissue
XX
XX Claim 27; Page 714-715; 824pp; English.
XX
XX AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for
XX human secreted proteins, and encode the proteins given in AAY12261 to
XX AAY12514, respectively. The proteins given represent the signal peptide
XX and an N-terminal fragment of a secreted protein. The nucleic acid
XX sequences can be used for producing secreted human gene products. They
XX can also be used to develop products for diagnosis and therapy. The
XX proteins obtained may have cytokine activity, cell
XX proliferation/differentiation activity, haematopoiesis regulating
XX activity, tissue growth regulating activity, reproductive hormone
XX regulating activity, chemotactic/chemokinetic activity, haemostatic and
XX thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX activity, tumour inhibition activity or other activities. The products
XX can be used in forensic, gene therapy and chromosome mapping procedures.
XX The sequences can also be used for obtaining corresponding promoter
XX sequences. The nucleic acids encoding the signal peptide can be used for
XX directing extracellular secretion of a polypeptide or the insertion of a
XX polypeptide into a membrane, or importing a polypeptide into a cell.
XX
XX Sequence 139 AA:
XX
XX Query Match 52.0%; Score 748; DB 20; Length 139;
XX Best Local Similarity 99.3%; Pred. No. 6.2e-70;
XX Matches 138; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 63 MAAAWSPGSPAPEAVTARLVGVLMFVSVTITGPGAVATAGGSESLKCEDLKVGVIC 120
DQ 1 MAAAWSPGSPAPEAVTARLVGVLMFVSVTITGPGAVATAGGSESLKCEDLKVGVIC 60
QY 123 PKINDATQEPVNCNTNTAHVSCFPAPNITCKDSSGNETHFTGNEVGFFKPISCRNVNG 180
DQ 61 PKINDATQEPVNCNTNTAHVSCFPAPNITCKDSSGNETHFTGNEVGFFKPISCRNVNG 120
QY 183 YKVAVALSLFLGLGADRF 201
DQ 121 YKVAVALSLFLGLGADRF 139

```

```
RESULT 5
AAY36021
ID AAY36021 standard; Protein: 162 AA.
XX AC AAY36021;
XX AC AAY36021;
XX DI 13-SEP-1999 (first entry)
XX DE Extended human secreted protein sequence, SEQ ID NO. 406.
XX DE
XX KW Secreted protein; human; cytokine; cellular proliferation; cell movement;
XX KW cellular differentiation; immune system regulator; anti-inflammatory;
XX KW haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
XX KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
XX KW genetic disease.
XX OS Homo sapiens.
XX PN WO9931236-A2.
XX PD 24-JUN-1999.
XX PF 17-DEC-1998; 98WO-IB02122.
XX PR 10-AUG-1998; 98US-0096116.
XX PR 17-DEC-1997; 97US-0069957.
XX PR 09-FEB-1998; 98US-0074121.
XX PR 13-APR-1998; 98US-0081563.
XX PA (GEST ) GENSET.
XX PI Bougueleret L, Duclert A, Dumas Milne Edwards J;
XX DR WPI; 1999-385906/32.
XX DR N-PSDB; AAX97705.
XX PT New isolated human secreted proteins
XX PS Claim 9; Page 346-347; 516pp; English.
XX CC This sequence is encoded by an extended human secreted protein coding
XX CC sequence of the invention. The secreted proteins can be used in treating
XX CC or controlling a variety of human conditions. The secreted proteins may
XX CC act as cytokines or may affect cellular proliferation or differentiation.
XX CC or may act as immune system regulators, haematopoiesis regulators, tissue
XX CC growth regulators, regulators of reproductive hormones or cell movement
XX CC or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
XX CC tumour inhibition activity. The DNAs can be used in forensic procedures
XX CC to identify individuals or in diagnostic procedures to identify
XX CC individuals having genetic diseases resulting from abnormal expression of
XX CC the genes corresponding to the extended cDNAs. They are also useful for
XX CC constructing a high resolution map of the human chromosomes. They can
XX CC also be used for gene therapy to control or treat genetic diseases.
XX SQ Sequence 162 AA;
Query Match 46.8%; Score 673.5; DB 20; Length 162;
Best Local Similarity 84.1%; Pred. No. 4.8e-62;
Matches 127; Conservative 4; Mismatches 17; Indels 3; Gaps 2;
QY 63 MAAWPSPGSAPEAVTARLVGLVFWVSVTTPGAVATSGAGSESLKCEDLVKGVYICKD 122
DB 1 MAAWPSPGSAPEAVTARLVGLVFWVSVTTPGAVATSGAGSESLKCEDLVKGVYICKD 60
QY 123 PKINDATQEPVNCNTNYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFKPISCRNVNGYS 182
DB 61 PKINDATQEPVNCNTNYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFKPISCRNVNGYS 120
QY 183 YKVAVALSLFLGMLGADRFLYLG--PALGLL 212
DB 121 YNEQSHVS--FSNMVGSRSILPWPCFGFVK 149
```

```
RESULT 6
AAY12426
ID AAY12426 standard; Protein: 148 AA.
XX AC AAY12426;
XX AC AAY12426;
XX DI 17-JUN-1999 (first entry)
XX DE Human 5' EST secreted protein SEQ ID NO:457.
XX DE
XX KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
XX KW forensic; gene therapy; chromosome mapping; signal peptide;
XX KW upstream regulatory sequence; cytokine activity; cell proliferation;
XX KW differentiation; haematopoiesis regulation; tissue growth regulation;
XX KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
XX KW thrombolytic; anti-inflammatory; tumour inhibition.
XX OS Homo sapiens.
XX PN WO9906548-A2.
XX PD 11-FEB-1999.
XX PF 31-JUL-1998; 98WO-IB01222.
XX PR 31-AUG-1997; 97US-0905135.
XX PA (GEST ) GENSET.
XX PI Duclert A, Dumas Milne Edwards J, Lacroix B;
XX DR WPI; 1999-153778/13.
XX DR N-PSDB; AAX41259.
XX PT New nucleic acids encoding human secreted proteins - obtained from
XX PT cDNA libraries prepared from e.g. liver, ovary, brain, prostate,
XX PT kidney, lung, umbilical cord, placenta and colon tissue
XX PS Claim 27; Page 763-764; 824pp; English.
XX CC AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for
XX CC human secreted proteins, and encode the proteins given in AAY12261 to
XX CC AAY12514, respectively. The proteins given represent the signal peptide
XX CC and an N-terminal fragment of a secreted protein. The nucleic acid
XX CC sequences can be used for producing secreted human gene products. They
XX CC can also be used to develop products for diagnosis and therapy. The
XX CC proteins obtained may have cytokine activity, cell
XX CC proliferation/differentiation activity, haematopoiesis regulating
XX CC activity, tissue growth regulating activity, reproductive hormone
XX CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
XX CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX CC activity, tumour inhibition activity or other activities. The products
XX CC can be used in forensic, gene therapy and chromosome mapping procedures.
XX CC The sequences can also be used for obtaining corresponding promoter
XX CC sequences. The nucleic acids encoding the signal peptide can be used for
XX CC directing extracellular secretion of a polypeptide or the insertion of a
XX CC polypeptide into a membrane, or importing a polypeptide into a cell.
XX SQ Sequence 148 AA;
Query Match 46.4%; Score 667.5; DB 20; Length 148;
Best Local Similarity 83.3%; Pred. No. 1.8e-61;
Matches 125; Conservative 4; Mismatches 18; Indels 3; Gaps 2;
QY 63 MAAWPSPGSAPEAVTARLVGLVFWVSVTTPGAVATSGAGSESLKCEDLVKGVYICKD 122
DB 1 MAAWPSPGSAPEAVTARLVGLVFWVSVTTPGAVATSGAGSESLKCEDLVKGVYICKD 60
QY 123 PKINDATQEPVNCNTNYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFKPISCRNVNGYS 182
DB 61 PKINDATQEPVNCNTNYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFKPISCRNVNGYS 120
QY 183 YKVAVALSLFLGMLGADRFLYLG--PALGLL 212
```


PD 28-NOV-2002.
 XX
 PF 19-MAR-2002; 2002WO-US09785.
 XX
 PR 21-MAR-2001; 2001US-277340P.
 PR 19-JUL-2001; 2001US-306171P.
 PR 13-NOV-2001; 2001US-331287P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2003-129429/12.
 XX
 PT Novel human secreted proteins, useful for detecting, preventing,
 PT diagnosing, prognosticating, treating and/or ameliorating
 PT cardiovascular disorders such as arrhythmia -
 XX
 PS Claim 13; SEQ ID 709; 1881pp; English.
 XX
 CC The present invention relates to novel human secreted proteins
 CC (AB47633-ABR48145) and their coding sequences (ACC50344-ACC50856). The
 CC proteins and their coding sequences are useful for the preparation of a
 CC diagnostic or pharmaceutical composition for diagnosing or treating a
 CC cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,
 CC coronary arteriosclerosis and myocardial ischemia), neural disorders,
 CC immune system disorders, muscular disorders, reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, renal disorders,
 CC proliferative disorders and/or cancerous diseases and conditions, for
 CC wound healing and epithelial cell proliferation, to treat inflammation or
 CC infection, for treating thrombosis and arteriosclerosis, for treating or
 CC preventing neural damage which occurs in neuronal disorders or
 CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's
 CC disease, to enhance bone and periodontal regeneration and aid in tissue
 CC transplants or bone grafts, to prevent skin aging or hair loss, to
 CC stimulate growth and differentiation of haematopoietic cells and bone
 CC marrow cells when used in combination with other cytokines, to maintain
 CC organs before transplantation or for supporting cell culture of primary
 CC tissues, to increase or decrease differentiation or proliferation of
 CC embryonic stem cells, or to modulate mammalian characteristics or
 CC metabolism.
 CC Note: the sequence data for this patent was published in electronic
 CC format and is available from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 221 AA;
 Query Match 14.0%; Score 201; EB 24; Length 221;
 Best Local Similarity 45.7%; Pred. No. 1.9e-12;
 Matches 53; Conservative 12; Mismatches 39; Indels 12; Gaps 5;
 QY 135 CTNYTA--HVSC----FPAPNITCKDSSGNE:HFTGNFVGFKPISCRNNGYSYKVAVA 183
 Db 112 CINSISCTVSC:PRORYPA-NCTVRD----HVHCLGNRT-FPKMLYCNKTCGYKASTALA 165
 QY 189 LSLFLGWLGAADRFYLGYPALGLLCKTFCGCGISGLIDFLISMQIVGSPGSSSYI 244
 Db 166 LSITLGGFGADRFLYGQWKEGLKFLSFGG:GIWTLIDVLLIGVGVGPADGSELYI 221
 RESULT 10
 ABR00112
 ID ABR00112 standard; Protein: 221 AA.
 XX
 AC ABR00112;
 XX
 DT 03-APR-2003 (first entry)
 XX
 DE Human gene 102 encoded secreted protein HMEED17, SEQ ID NO:401.
 XX
 KW Human; secreted protein; digestive disorder; gastrointestinal disorder;
 KW mouth; oesophagus; stomach; small intestine; large intestine; liver;
 KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;

KW immune disorder; inflammation; infection; wound healing; drug screening;
 KW chromosome identification; chromosome mapping; cytostatic;
 KW antiinflammatory; immunosuppressive; vulnery; gene therapy.
 XX
 OS Homo sapiens.
 OS WO200276488-A1.
 PN
 XX
 PD 03-OCT-2002.
 XX
 PF 19-MAR-2002; 2002WO-US08276.
 XX
 PR 21-MAR-2001; 2001US-277340P.
 PR 19-JUL-2001; 2001US-306171P.
 PR 13-NOV-2001; 2001US-331287P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2003-029900/02.
 DR N-PSDB; ABZ71291.
 XX
 PT New human secreted proteins and nucleic acids, useful for detecting,
 PT preventing, diagnosing, prognosticating, treating and/or ameliorating
 PT e.g. gastrointestinal diseases and disorders, or cancers -
 XX
 PS Claim 13; Page 1007; 1216pp; English.
 XX
 CC ABZ71190-ABZ71478 represent cDNAs corresponding to 178 human secreted
 CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.
 CC ABZ71479-ABZ71540 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins
 CC the use of the secreted proteins in drug screening, and recombinant
 CC vectors and host cells comprising a nucleic acid of the invention. The
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody
 CC fragments specific for the secreted proteins, and modulators of protein
 CC activity are useful for diagnosing, treating, ameliorating or preventing
 CC digestive disorders. Such conditions include disorders of the mouth,
 CC oesophagus, stomach, small intestine, large intestine, liver, biliary
 CC tract and pancreas, and include cancers of these organs and tissues. The
 CC secreted proteins and their nucleic acids may also be used in the
 CC treatment of immune disorders, inflammation, infection,
 CC hyperproliferative disorders, and to promote wound healing. Nucleic acids
 CC of the invention may be used for chromosome identification, chromosome
 CC mapping, in gene therapy, for identifying individuals from minute
 CC biological samples, as hybridisation probes, and as molecular weight
 CC markers. The present sequence represents a human secreted protein of the
 CC invention.
 XX
 SQ Sequence 221 AA;
 Query Match 14.0%; Score 201; DB 24; Length 221;
 Best Local Similarity 45.7%; Pred. No. 1.9e-12;
 Matches 53; Conservative 12; Mismatches 39; Indels 12; Gaps 5;
 QY 135 CTNYTA--HVSC----FPAPNITCKDSSGNE:HFTGNFVGFKPISCRNNGYSYKVAVA 183
 Db 112 CINSISCTVSC:PRORYPA-NCTVRD----HVHCLGNRT-FPKMLYCNKTCGYKASTALA 165
 QY 189 LSLFLGWLGAADRFYLGYPALGLLCKTFCGCGISGLIDFLISMQIVGSPGSSSYI 244
 Db 166 LSITLGGFGADRFLYGQWKEGLKFLSFGG:GIWTLIDVLLIGVGVGPADGSELYI 221
 RESULT 11
 AAY27658
 ID AAY27658 standard; Protein: 222 AA.
 XX
 AC AAY27658;
 XX
 DT 30-JUL-1999 (first entry)
 XX

Human secreted protein encoded by gene No. 92.

DE Human: secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

XX Homo sapiens.

XX WO9924836-A1.

XX 20-MAY-1999.

XX 04-NOV-1998; 98WO-US23435.

XX 17-NOV-1997; 97US-0066100.

XX 07-NOV-1997; 97US-0064800.

XX 07-NOV-1997; 97US-0064308.

XX 07-NOV-1997; 97US-0064911.

XX 07-NOV-1997; 97US-0064912.

XX 07-NOV-1997; 97US-0064983.

XX 07-NOV-1997; 97US-0064984.

XX 07-NOV-1997; 97US-0064985.

XX 07-NOV-1997; 97US-0064987.

XX 07-NOV-1997; 97US-0064988.

XX 17-NOV-1997; 97US-0066090.

XX 17-NOV-1997; 97US-0066094.

XX 17-NOV-1997; 97US-0066095.

XX 17-NOV-1997; 97US-0066099.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Carter KC, Ebner R, Endress GA, Feng P, Janat F;

XX Kyaw H, Lafleur DW, Moore PA, Ni J, Olsen HS, Rosen CA;

XX Ruben SM, Shi Y, Soppet DR, Wei Y;

XX WPI: 1999-337740/28.

XX N-PSDB: AAX85024.

XX New human secreted proteins and coding sequences useful for treating

XX disorders of the immune system and hyperproliferative disorders

XX Claim 1; Page 400-401; 507pp; English.

XX This sequence represents a secreted human protein encoded by the gene

XX clone detailed in the descriptor line. The gene can be used to generate

XX fusion proteins by linking to the gene to a human immunoglobulin FC

XX portion (e.g. AAX84924) for increasing the stability of the fused

XX protein as compared to the human protein only.

XX The invention relates to 125 novel genes and their fragments (nucleic

XX acid sequences: AAX84933-X85057; amino acid sequences AAY27567-Y27933;

XX which are useful for preventing, treating or ameliorating medical

XX conditions e.g. by protein or gene therapy. Also, pathological

XX conditions can be diagnosed by determining the amount of the new

XX polypeptides in a sample or by determining the presence of mutations in

XX the new polynucleotides. Specific issues are described for each of the 125

XX polynucleotides, based on which tissues they are most highly expressed in

XX (see AAX84933 for described uses).

XX Sequence 222 AA;

XX Query Match 14.0%; Score 201; DB 20; Length 222;

XX Best Local Similarity 45.7%; Pred. No. 1.9e-12;

XX Matches 53; Conservative 12; Mismatches 39; Indels 12; Gaps 5;

QY 135 CTNYTA--HVSC-----FPAPNITCKDSSGNETHFTGNEVGFPPICSRNVNGYSKYAVA 188

DB 112 CTNSTSCMTVSCPQRYPN-NCTVRD----HVHCLGNRT-FPKMLYCNMTGGYKWTALA 165

QY 189 LSLFLGWLGADEYLYGYPALGSLKFTCTVGCFGIGSLGIFLISMQIVGSPDGSYI 244

DB 166 LSLTGGFADRYLQGWXESGKLFSGGLGINTLIDVLLGVGVGVPADGSLYI 221

RESULT 12

AAY70761

ID AAY70761 standard: Protein; 221 AA.

XX AAY70761;

XX 24-JUL-2000 (first entry)

XX Human beta-amyloid peptide (BAP) binding protein, BBP3.

XX Beta-amyloid peptide binding protein; BBP; BAP; tumour; suppressor;
 KW G-protein coupled receptor; GPCR; integral membrane protein; antigen;
 KW neuronal cell; nonhuman primate; NHP; G-protein signalling pathway;
 KW apoptosis; immunogen; therapeutic; treatment; prevention; diagnostic.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Domain 154..175

XX Domain /label= Transmembrane_domain_1

XX Domain 176..178

XX Domain /label= DRE_motif

XX Domain /note= Substitution of the Arg abrogates protection"

XX Domain 190..215

XX Domain /label= Transmembrane_domain_2

XX WO200022125-A2.

XX 20-APR-2000.

XX 13-OCT-1999; 99WO-US21621.

XX 13-OCT-1998; 98US-0104104.

XX (AMHP) AMERICAN HOME PROD CORP.

XX Ozenberger BA, Kajkowski EM, Io CF;

XX WPI: 2000-317982/27.

XX N-PSDB: AAY52371.

XX Novel G-protein-coupled receptor-like proteins and polynucleotides
 XX useful for regulating apoptosis, comprises integral membrane protein
 XX traversing the membrane twice

XX Example 1; Page 67-68; 68pp; English.

XX The present sequence is the beta-amyloid peptide (BAP) binding protein;
 XX (BBP3). It is an integral membrane protein, that traverse the membrane
 XX twice. It is related to G protein-coupled receptor (GPCR) protein
 XX superfamily. It interacts with G-alpha proteins and regulates the
 XX activity of G-protein signalling pathways. BAP genes are widely expressed
 XX in neuronal cells of nonhuman primate (NHP) brain and overexpressed in
 XX some tumours. It functions as a suppressor of apoptosis induction. BAP
 XX proteins are used as immunogens to raise antibodies, useful as
 XX therapeutics and as antigens in solid phase assays. They are also useful
 XX as reagents to identify molecules which effect the interaction of BBP and
 XX a cloned protein, that are useful in the treatment or prevention of
 XX diseases associated with apoptosis. The polynucleotides are useful for
 XX diagnostics.

XX Sequence 221 AA;

XX Query Match 13.9%; Score 200; DB 21; Length 221;

XX Best Local Similarity 45.7%; Pred. No. 2.4e-12;

XX Matches 53; Conservative 12; Mismatches 39; Indels 12; Gaps 5;

QY 135 CTNYTA--HVSC-----FPAPNITCKDSSGNETHFTGNEVGFPPICSRNVNGYSKYAVA 188

DB 112 CTNSTSCMTVSCPQRYPN-NCTVRD----HVHCLGNRT-FPKMLYCNMTGGYKWTALA 165

[illegible]

RESULT 13

ABR48493	ID	ABR48493	Standard: Protein: 221 AA.
XX	AC	AC	
XX	ABR48493:		
XX		13-JUN-2003	(first entry)
XX	DI	Human Amyloid Apoptotic Receptor (AAR; protein.	
XX	DE		
XX	XX		
XX	KW	Human: GENSET; therapeutic; therapy.	

```

Query Match      13.9%  Score 200;  DB 24;  Length 221;
Best Local Similarity 45.7%;  Pred. No. 2.4e-12;
Matches 53;  Conservative 12;  Mismatches 39;  Indels 12;  Caps 5;

QY 135 CTNYTA--HVSC-----FPAPNITCKUSSGNETHFTGNEVGFFKPISCRNNVNGYSKYKAVA 148
      ||| |  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 112 CTNSTSMTVSCPQRKYPA-NCTVRD----HVHCLGNRT-FPKMLYCNWTGGYKWSATA 165
      ||| |  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 189 LSLFLGWLGNDRYLVGPA-GLLKFCTVCPCCGLSLIDFILISMCIVGPPSGSGSY 244
      |||  |||  ||||| |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 166 LSTLGGFGADRYLVGWRREGLGKLSFSGIGWTLLDVLIIYGVYGPAGGSHV 221
      |||  |||  ||||| |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

RESULT 14
AAB08641

ID	AA08641	standard; Protein; 247 AA.
XX	AA08641	
AC	AA08641	
XX	AA08641	
DI	02-JAN-2001	(first entry)
XX	02-JAN-2001	(first entry)
DE	A human TANGO 223 polypeptide.	
XX	A human TANGO 223 polypeptide.	
XX	TANGO 201; TANGO 223; secreted protein; transmembrane protein; pro-inflammatory disorder; neoplasm; tumour; pancreatic disorder; cirrhosis; pancreatitis; thyroiditis; goiter; Graves' disease; gastric disorder; gastritis; tumour; placental disorder; placentitis; spontaneous abortion; pulmonary disorder; atelectasis; edema; Goodpasture's syndrome; muscular dystrophy; cardiovascular disorder; ischemic heart disease; congenital heart disease; cerebral edema; cerebrovascular disease; hepatic disorder; jaundice; hepatitis; leukocytic disorder; hyperplasia; leukopenia; leukocytosis; malignant lymphoma; prostate disorder; inflammatory disease.	
XX	TANGO 201; TANGO 223; secreted protein; transmembrane protein; pro-inflammatory disorder; neoplasm; tumour; pancreatic disorder; cirrhosis; pancreatitis; thyroiditis; goiter; Graves' disease; gastric disorder; gastritis; tumour; placental disorder; placentitis; spontaneous abortion; pulmonary disorder; atelectasis; edema; Goodpasture's syndrome; muscular dystrophy; cardiovascular disorder; ischemic heart disease; congenital heart disease; cerebral edema; cerebrovascular disease; hepatic disorder; jaundice; hepatitis; leukocytic disorder; hyperplasia; leukopenia; leukocytosis; malignant lymphoma; prostate disorder; inflammatory disease.	
XX	TANGO 201; TANGO 223; secreted protein; transmembrane protein; pro-inflammatory disorder; neoplasm; tumour; pancreatic disorder; cirrhosis; pancreatitis; thyroiditis; goiter; Graves' disease; gastric disorder; gastritis; tumour; placental disorder; placentitis; spontaneous abortion; pulmonary disorder; atelectasis; edema; Goodpasture's syndrome; muscular dystrophy; cardiovascular disorder; ischemic heart disease; congenital heart disease; cerebral edema; cerebrovascular disease; hepatic disorder; jaundice; hepatitis; leukocytic disorder; hyperplasia; leukopenia; leukocytosis; malignant lymphoma; prostate disorder; inflammatory disease.	
OS	Homo sapiens.	
XX	Homo sapiens.	
XX	Homo sapiens.	
FT	Key	Location/Qualifiers
FT	Peptide	1..29
FT	Modified-site	/note= "signal peptide"
FT	Modified-site	28..31
FT	Modified-site	/note= "casein kinase II phosphorylation site"
FT	Modified-site	29..34
FT	Modified-site	/note= "myristoylation site"
FT	Protein	30..247
FT	Domain	/note= "mature protein"
FT	Domain	30..215
FT	Domain	/note= "extracellular domain"
FT	Modified-site	39..42
FT	Modified-site	/note= "casein kinase II phosphorylation site"
FT	Modified-site	39..41
FT	Modified-site	/note= "protein kinase C phosphorylation site"
FT	Modified-site	97..90
FT	Modified-site	/note= "N-glycosylation site"
FT	Modified-site	107..110
FT	Modified-site	/note= "casein kinase II phosphorylation site"
FT	Modified-site	115..117
FT	Modified-site	/note= "protein kinase C phosphorylation site"
FT	Modified-site	122..125
FT	Modified-site	/note= "N-glycosylation site"
FT	Modified-site	124..126
FT	Modified-site	/note= "protein kinase C phosphorylation site"
FT	Modified-site	134..137
FT	Modified-site	/note= "casein kinase II phosphorylation site"
FT	Modified-site	140..143
FT	Modified-site	/note= "N-glycosylation site"
FT	Modified-site	157..160
FT	Modified-site	/note= "N-glycosylation site"
FT	Modified-site	159..161
FT	Modified-site	/note= "protein kinase C phosphorylation site"
FT	Modified-site	159..162
FT	Modified-site	/note= "casein kinase II phosphorylation site"
FT	Modified-site	169..172
FT	Modified-site	/note= "N-glycosylation site"
FT	Modified-site	179..182
FT	Modified-site	/note= "N-glycosylation site"
FT	Modified-site	197..202
FT	Modified-site	/note= "myristoylation site"
FT	Domain	216..238
FT	Domain	/note= "transmembrane domain"
FT	Modified-site	226..229
FT	Modified-site	/note= "casein kinase II phosphorylation site"
FT	Domain	239..247
FT	Domain	/note= "cytoplasmic domain"
XX	WO2006050442-A2.	
XX	31-AUG-2000.	
XX	31-AUG-2000.	
XX	31-AUG-2000.	

```

PF 25-FEB-2000; 2000WO-US04784.
XX
XX
PR 26-FEB-1999; 99US-0259388.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX McCarthy SA;
XX
XX WPI: 2000-572066/53.
DR N-PSDB; AAA64408, AAA64409.
XX
XX Novel nucleic acid molecule encoding secreted or transmembrane protein
PT useful for identifying modulators and for diagnosing and treating
PT pancreatic, cardiovascular, liver and pituitary disorders -
XX
XX Claim 9; Fig 8; 176pp; English.
XX
XX The present sequence represents TANGO 223 polypeptide. The specification
CC also describes a TANGO 201 polypeptide. These polypeptides are secreted
CC or transmembrane proteins. Human TANGO 201 and 223 nucleic acids,
CC proteins and their modulators are useful for treating proliferative
CC disorders e.g. neoplasms or tumours, pancreatic disorders
CC (e.g. pancreatitis), disorders of the adrenal cortex, adrenal medulla,
CC thyroid gland (e.g. thyroiditis), goiter, Graves' disease, gastric
CC disorders (e.g. gastritis or tumours), placental disorders (e.g.
CC placentalis or spontaneous abortion), pulmonary disorders (e.g.
CC atelectasis), edema, Goodpasture's syndrome, disorders of the skeletal
CC muscle (e.g. muscular dystrophy), cardiovascular disorders (e.g.
CC ischemic heart disease and congenital heart disease), disorders of the
CC brain (e.g. cerebral edema), cerebrovascular disease and to treat
CC injury or trauma to the brain. They are also useful for treating
CC hepatic disorders (e.g. jaundice, hepatitis, cirrhosis or malignant
CC tumours), renal, testicular, intestinal disorders. TANGO 223
CC polynucleotides are also useful for treating leukocytic disorders (e.g.
CC leukopenias, leukocytosis and malignant lymphomas) and prostate
CC disorders (e.g. inflammatory diseases, hyperplasia or tumours).
XX
XX Sequence 247 AA:
Query Match 13.9%; Score 200; DB 21; Length 247;
Best Local Similarity 45.7%; Pred. No. 2.8e-12;
Matches 53; Conservative 12; Mismatches 39; Indels 12; Gaps 5;

QY 135 CTNYTA--HVSC---FPAPNITCKDSSGNETHFTGNEVGFEPKIPISCRNVNGYSYKVAVA 188
11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366 1367 1368 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392 1393 1394 1395 1396 1397 1398 1399 1400 1401 1402 1403 1404 1405 1406 1407 1408 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419 1420 1421 1422 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488 1489 1490 1491 1492 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502 1503 1504 1505 1506 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516 1517 1518 1519 1520 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530 1531 1532 1533 1534 1535 1536 1537 1538 1539 1540 1541 1542 1543 1544 1545 1546 1547 1548 1549 1550 1551 1552 1553 1554 1555 1556 1557 1558 1559 1560 1561 1562 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572 1573 1574 1575 1576 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586 1587 1588 1589 1590 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600 1601 1602 1603 1604 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614 1615 1616 1617 1618 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628 1629 1630 1631 1632 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642 1643 1644 1645 1646 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656 1657 1658 1659 1660 1661 1662 1663 1664 1665 1666 1667 1668 1669 1670 1671 1672 1673 1674 1675 1676 1677 1678 1679 1680 1681 1682 1683 1684 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696 1697 1698 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709 1710 1711 1712 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723 1724 1725 1726 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751 1752 1753 1754 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765 1766 1767 1768 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779 1780 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793 1794 1795 1796 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807 1808 1809 1810 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821 1822 1823 1824 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835 1836 1837 1838 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849 1850 1851 1852 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863 1864 1865 1866 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877 1878 1879 1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891 1892 1893 1894 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 25
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw mode:

Run on: September 26, 2003, 17:25:22 : Search time 30 seconds
(without alignments)
379,387 Million cell updates/sec

Title: US-09-852-100A-2

Perfect score: 1439

Sequence: 1 MHILKGSPNVIPRAHGQKNI.....TRLRLSINETFRKIQIYP 269

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2.6/ptodata/1/1aa/5A-COMB.pep.*
- 2: /cgn2.6/ptodata/1/1aa/5B-COMB.pep.*
- 3: /cgn2.6/ptodata/1/1aa/6A-COMB.pep.*
- 4: /cgn2.6/ptodata/1/1aa/6B-COMB.pep.*
- 5: /cgn2.6/ptodata/1/1aa/PCTUS-COMB.pep.*
- 6: /cgn2.6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed. and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85	5.9	310	2	US-08-414-657D-45 Sequence 45, Appl
2	85	5.9	338	2	US-08-414-657D-42 Sequence 42, Appl
3	85	5.9	338	2	US-08-414-657D-43 Sequence 43, Appl
4	85	5.9	338	4	US-09-135-080-4 Sequence 4, Appl
5	84	5.8	258	4	US-09-328-352-4253 Sequence 4253, Ap
6	83	5.8	764	2	US-08-177-109A-2 Sequence 2, Appl
7	83	5.8	764	2	US-08-687-706-2 Sequence 2, Appl
8	81	5.6	797	3	US-09-182-728A-2 Sequence 2, Appl
9	81	5.6	797	4	US-09-795-232-2 Sequence 2, Appl
10	80.5	5.6	150	4	US-09-252-991A-16958 Sequence 16958, A
11	80.5	5.6	304	2	US-08-414-657D-44 Sequence 44, Appl
12	80.5	5.6	325	2	US-08-414-657D-2 Sequence 2, Appl
13	80.5	5.6	325	2	US-08-414-657D-41 Sequence 41, Appl
14	80.5	5.6	325	4	US-09-135-080-2 Sequence 2, Appl
15	80.5	5.6	456	3	US-09-058-389A-4 Sequence 4, Appl
16	80.5	5.6	456	4	US-09-611-781-4 Sequence 4, Appl
17	80.5	5.6	1237	1	US-08-241-853-2 Sequence 2, Appl
18	80.5	5.6	1237	4	US-08-850-917-2 Sequence 2, Appl
19	80	5.6	328	4	US-09-253-316-28 Sequence 28, Appl
20	80	5.6	660	3	US-09-181-706-8 Sequence 8, Appl
21	80	5.6	660	3	US-09-458-791-8 Sequence 8, Appl
22	80	5.6	660	3	US-09-459-066-8 Sequence 8, Appl
23	80	5.6	660	4	US-09-459-065-8 Sequence 8, Appl
24	79.5	5.5	354	2	US-08-700-013B-9 Sequence 9, Appl
25	79.5	5.5	797	2	US-08-700-013B-21 Sequence 21, Appl
26	79.5	5.5	927	4	US-09-328-352-7922 Sequence 7922, Ap
27	79	5.5	797	2	US-08-700-013B-19 Sequence 19, Appl

Sequence 124, Appl
Sequence 19777, A
Sequence 60, Appl
Sequence 8, Appl
Sequence 2, Appl
Sequence 31718, A
Sequence 26357, A
Sequence 120, Appl
Sequence 122, Appl
Sequence 4, Appl
Sequence 26323, A
Sequence 2, Appl
Sequence 2, Appl
Sequence 21426, A
Sequence 3, Appl
Sequence 37, Appl
Sequence 11, Appl
Sequence 42, Appl

ALIGNMENTS

RESULT 1
US-08-414-657D-45
: Sequence 45, Application US/08414657D
: Patent No. 5861283
: GENERAL INFORMATION:
: APPLICANT: Iovitt, Pat
: APPLICANT: Pimenta, Aurea
: APPLICANT: Fischer, Itzhak
: APPLICANT: Zhukareva, Victoria
: TITLE OF INVENTION: Limbic System-Associated Membrane
: TITLE OF INVENTION: Protein and DNA
: NUMBER OF SEQUENCES: 60
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dechert Price & Rhoads
: STREET: 997 Lenox Drive, Building 3, Suite 210
: CITY: Lawrenceville
: STATE: NJ
: COUNTRY: USA
: ZIP: 08543
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/414.657D
: FILING DATE: 31-MAR-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Bloom, Allen
: REGISTRATION NUMBER: 29,135
: REFERENCE/DOCKET NUMBER: 317743-102
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 609-520-3214
: TELEFAX: 609-520-3259
: TELEX:
: INFORMATION FOR SEQ ID NO: 45:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 310 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear

US-08-414-657D-45

Query Match 5.9%; Score 85; DB 2; Length 310;
Best Local Similarity 27.7%; Pred. No. 1.7;
Matches 36; Conservative 15; Mismatches 47; Indels 32; Gaps 7;

```

01 101 SAGGESLKCEDKVG-----QYICKDPKINDATQEPVNCNTNYTAHVSCFPAPNITCKDSS 154
02 102 TTGROASLKCEASAVPAPDFEWYRDDTRINSANGLEIKSTE--GOSSLTVTNTVT--EEHY 257
03 157 GNEHFTTNEVG-----FFKPITSICRVNVGYSYKVAVALISFLGWLGADRFFYGYPALG 209
04 258 GNYTCVAANKILGVNNSLVLPFGSVRGNG--SISLAVPL-----WL-----LA 300
05 210 LKCTGVGFC 219
06 30: ASIFCLLSKC 310
07
08
09
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1
```

```

QY      210 LLKFTVGF 219
      || : |
Db      329 ASLFCLLSK 338

RESULT 3
US-08-414-657D-43
: Sequence 43, Application US/08414657D
: Patent No. 5861283
: GENERAL INFORMATION:
: APPLICANT: Levitt, Pat
: APPLICANT: Pimental, Aurea
: APPLICANT: Fischer, Itzhak
: APPLICANT: Zhukareva, Victoria
: TITLE OF INVENTION: Limbic System-Associated Membrane
: TITLE OF INVENTION: Protein and DNA
: NUMBER OF SEQUENCES: 60
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dechert Price & Rhoads
: STREET: 997 Lenox Drive, Building 3, Suite 210
: CITY: Lawrenceville
: STATE: NJ
: COUNTRY: USA
: ZIP: 08543
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/414,657D
: FILING DATE: 31-MAR-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Bloom, Allen
: REGISTRATION NUMBER: 29,135
: REFERENCE/DOCKET NUMBER: 317743-102
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 609-520-3214
: TELEFAX: 609-520-3259
: TELEX:
: INFORMATION FOR SEQ ID NO: 43:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 338 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-414-657D-43

Query Match 5.9%; Score 85; DB 2; Length 338;
Best Local Similarity 27.7%; Pred. No. 2;
Matches 36; Conservative 15; Mismatches 47; Indels 32; Gaps 7;

QY      101 SAGESLKDLDKVG-----QYCKDKPKINDATQEPVNCNTYTAHVSCFPAPNITCKDSS 156
Db      230 ITGRQASLKGCEASVAPDPDEWFRDDTRINSANGLEIKSTE---GQSLSLTVNTVTEHY 285

QY      157 GNTHETGTGNEVG-----FFKPLSCRNNGSYKVAVALSLFLGLWLGADRFYIGYPALG 209
Db      286 GNTCTVAANKLGYTNASLVLFRCGSVRGNG-SISLAVPL-----WL-----LA 328

QY      210 LLKFTVGF 219
      || : |
Db      329 ASLFCLLSK 338

RESULT 4
US-09-135-080-4
: Sequence 4, Application US/09135080
: Patent No. 6423827

```



```

; GENERAL INFORMATION:
; APPLICANT: Levitt, Pat R.
; ORGANISM: Pimenta, Aurea
; APPLICANT: Fischer, Itzhak
; APPLICANT: Zhukareva, Victoria
; TITLE OF INVENTION: Limbic System-Associated Membrane
; TITLE OF INVENTION: Protein and DNA
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 997 Lenox Drive, Building 3, Suite 210
; CITY: Lawrenceville
; STATE: NJ
; COUNTRY: USA
; ZIP: 08543
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/135,080
; FILING DATE: 17-AUG-1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/414,657
; FILING DATE: 31-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bloom, Allen
; REGISTRATION NUMBER: 29,135
; REFERENCE/DOCKET NUMBER: 317743-102A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 609-620-3214
; TELEFAX: 609-620-3259
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 338 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-135-080-4

Query Match 5.9%; Score 85; DB 4; Length 338;
Best Local Similarity 27.7%; Pred. No. 2;
Matches 36; Conservative 15; Mismatches 47; Indels 32; Gaps 7;

QY 101 SAGHESLACEDLKVG-----QYICKDKPKINDATCEPVNCTNYTAHVSCEPPAPNITCKDS 156
DB 230 TGRQASLKCEASAVPADEFWYKDDIRNSANGLEIKS:E---GSSSLVTNVI EEHY 285
QY 157 GNETHFTGNEVG-----FFKPSICRNNGYSYKVAVALSIFLGLGADRFYLGYPALG 204
DB 286 GNYTCVANKLGYTNASLVLFRCGSVRGNG-S:SLAVPL-----WL-----LA 326
QY 210 LLKPFCTVGF 219
DB 329 ASLFCLLSKC 338

RESULT 5
US-09-328-352-4253
; Sequence 4253, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4253

```

```

; LENGTH: 258
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-4253

Query Match 5.8%; Score 84; DB 4; Length 258;
Best Local Similarity 23.5%; Pred. No. 1.7;
Matches 52; Conservative 24; Mismatches 93; Indels 52; Gaps 10;

QY 78 TARLVGVL---WFSVVTTPW-----GAVATSAGSESLKCEDLVK-----G 116
DB 57 TTGLYGLPLNVEWITRLERGPYWSEKIDKGTFFRGAPGSISIRSPDYPSPISQPAADGG 116
QY 117 QYICKDKPKINDATCEPVNCTNY--TAHVSC-PPANITCKDSSGNETHFTGNEVGFFKPI 173
DB 117 FYLDKDPK-----EPVKIYRFTTKAVPVEVPSDNVTC-----NTLAYTRP 158
QY 174 SCRNNGYSYKVAVALSFLGLGADRFYLGYPALGCLLKFCTVGFQIGSLIDFILLISMQ 213
DB 159 ASHKVSLVSFATAGTVGGVTGALIGKNFSSGNKSYQATGAGAGGAIGGLIVAAIINAE 218
QY 234 ---IVG---PSDGSYIIDYVGTRLRLSLTNETFRKIQLYP 269
DB 219 VGKIIGGLPIKESPM-----EKLRELGAKEPLKQISLLP 254

RESULT 6
US-08-177-109A-2
; Sequence 2, Application US/08177109A
; Patent No. 5869615
; GENERAL INFORMATION:
; APPLICANT: Dennis E. Hourcade and Teresa J. Oglesby
; TITLE OF INVENTION: MODIFIED COMPLEMENT PROTEASES
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/177,109A
; FILING DATE: 03-JAN-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,264
; REFERENCE/DOCKET NUMBER: WU 107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 764 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-177-109A-2

Query Match 5.8%; Score 83; DB 2; Length 764;
Best Local Similarity 24.1%; Pred. No. 10;
Matches 49; Conservative 21; Mismatches 71; Indels 62; Gaps 12;

QY 24 GTGLYPMRGPFKNLALLPFLPLPLGG-----GSGSGEKVSV----- 60

```

Db 2 GSNLSP-----QCLMPPFILGLSGGVTTPWSLAQPQCSLEGVIEIKGSPFLQEG 55
Qy 61 SKMAAMPSSG--PSAPEAVTARLVGLVNFVSVITGPMCAVATS-----AGGESLKC-- 110
Db 56 QALEYVCSGPGYPVQVTRCR-----STGWSSTLKTDQKTVRAECRAHICPR 105
Qy 111 -EDLKVGQYICKPKINDATQEPVNC-TNYTAHVSCFPAPNITCKDSS--GNETHFTGNE 166
Db 106 PHDFENGWYPRSPYVNVSDRISFHCYDGYTLRGA-----NRTQVNGRWSGQTALCDNG 161
Qy 167 VGPFK-----PISCRNVNGYSYKV 185
Db 162 AGYCSNPGIPIGTRKV-GSQYRL 183

RESULT 7

US-08-687-706-2
: Sequence 2, Application US/04687706
: Patent No. 5928892
: GENERAL INFORMATION:
: APPLICANT: Dennis E. Hourcade and Teresa J. Oakesby
: TITLE OF INVENTION: MODIFIED COMPLEMENT PROTEASES
: NUMBER OF SEQUENCES: 62
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Patrea L. Pabst
: STREET: 2800 One Atlantic Center
: CITY: Atlanta
: STATE: Georgia
: COUNTRY: USA
: ZIP: 30309-3450
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/687,706
: FILING DATE: 26-JUL-1996
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/177,109
: FILING DATE: 03-JAN-1994
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Pabst, Patrea L.
: REGISTRATION NUMBER: 31,284
: REFERENCE/DOCKET NUMBER: WU 107 DIV
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (404) 873-8794
: TELEFAX: (404) 873-8795
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 764 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHEZICAL: NO
US-08-687-706-2

Query Match 5.8%; Score 83; DB 2; Length 764;
Best Local Similarity 24.1%; Pred. No. 10;
Matches 49; Conservative 21; Mismatches 7; Indels 62; Gaps 12;
Qy 24 GTGLYPMRGPKNALPFLPLGGG-----GSGSGFKSV----- 60
Db 2 GSNLSP-----QCLMPPFILGLSGGVTTPWSLAQPQCSLEGVIEIKGSPFLQEG 55
Qy 61 SKMAAMPSSG--PSAPEAVTARLVGLVNFVSVITGPMCAVATS-----AGGESLKC-- 110
Db 56 QALEYVCSGPGYPVQVTRCR-----STGWSSTLKTDQKTVRAECRAHICPR 105
Qy 111 -EDLKVGQYICKPKINDATQEPVNC-TNYTAHVSCFPAPNITCKDSS--GNETHFTGNE 166

Db 106 PHDFENGWYPRSPYVNVSDRISFHCYDGYTLRGA-----NRTQVNGRWSGQTALCDNG 161
Qy 167 VGPFK-----PISCRNVNGYSYKV 185
Db 162 AGYCSNPGIPIGTRKV-GSQYRL 183

RESULT 8
US-09-182-728A-2
: Sequence 2, Application US/09182728A
: Patent No. 623883
: GENERAL INFORMATION:
: APPLICANT: BROWN, ANTHONY
: APPLICANT: CHAPMAN, CONRAD GERALD
: APPLICANT: GLOGER, ISRAEL SIMON
: APPLICANT: EVANS, JOANNE RACHEL
: APPLICANT: CAIRNS, WILLIAM
: APPLICANT: HERDON, HUGH
: TITLE OF INVENTION: NOVEL COMPOUNDS
: FILE REFERENCE: GP-30176
: CURRENT APPLICATION NUMBER: US/09/182,728A
: CURRENT FILING DATE: 1998-10-29
: PRIOR APPLICATION NUMBER: 9818890.7
: PRIOR FILING DATE: 1998-08-28
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 2
: LENGTH: 797
: TYPE: PRT
: ORGANISM: HOMO SAPIENS
US-09-182-728A-2

Query Match 5.6%; Score 81; DB 3; Length 797;
Best Local Similarity 23.9%; Pred. No. 17;
Matches 47; Conservative 32; Mismatches 80; Indels 38; Gaps 21;
Qy 87 FSVTTGPMGAVATSAGGEESLKCD---LKVQGYICKD-PKINDATQEPVNCNTYAHV 442
Db 302 FVSVL--PMGSCNNPWTPE---CKDKTKLLDSCVISDHPKI-----QIKNSTFCM 348
Qy 143 SCPPAPNITCKDSSGNETHFTGNEVGFEPKIPISCRNVNGYSY----KVAVALSFJLWLGA 198
Db 349 TAYPNVTMNTSOANKTFVSGSE-EYFKYVLKISAGIEYPOEIRWPLALCLFLAW-- 405
Qy 199 DREYLGYPAL-----GLLKFCYGVGCGIGSLIDFILISMQIVGSPSGSSYLIDYVTR 251
Db 406 ----IVYASLAKGIK*SGKVYFTATFPYV-VLVILLIRGVTLPGAGAGIWFITPKWEK 460
Qy 252 LIRLSITNEIFRKTOLY 268
Db 461 LTNATVWKDA--ATOIF 475

RESULT 9

US-09-795-232-2
: Sequence 2, Application US/09795232
: Patent No. 6426405
: GENERAL INFORMATION:
: APPLICANT: Anthony M. Brown
: APPLICANT: Conrad Gerald Chapman
: APPLICANT: Israel Simon Gloger
: APPLICANT: Joanne Rachel Evans
: APPLICANT: William Cairns
: APPLICANT: Hugh Jonathan Herdon
: TITLE OF INVENTION: NOVEL COMPOUNDS
: FILE REFERENCE: GP-30176-D1
: CURRENT APPLICATION NUMBER: US/09/795,232
: CURRENT FILING DATE: 2001-02-28
: PRIOR APPLICATION NUMBER: 09/182,728
: PRIOR FILING DATE: 1998-10-29
: PRIOR APPLICATION NUMBER: 9818890.7
: PRIOR FILING DATE: 1998-08-28

```
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 797
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-795-232-2

Query Match.          5.6%; Score 81; DB 4; Length 797;
Best Local Similarity 23.9%; Pred. No. 17;
Matches 47; Conservative 32; Mismatches 80; Indels 38; Gaps 11;

QY 87 FVSVTTGPGCAVATISAGGESLACED--LKVGQYICKD-PKINDATQEPVNCNTYTAHV 142
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 302 FVSVL--PMGSCNPNWTP--CKDKTKLLSCVISHPKI-----QTKNSTFCM 348

QY 143 SCFPAPNITCKSSGNETHTFTGNEVGFFPKISCRNVNGYSY---KVAVALSFLGLWGA 196
: || : || : || : || : || : || : || : || : || : || : || : ||
Db 349 TAYPNVTWVNTSQANKTFVSGSE-EYFKYFVLKISAGIEYPGEIRWPLALCIFLAWV-- 405

QY 199 DRYLGYPAL-----GLLKFTVGFPGGIGSLIDFILISKQIVGSDGSSYLDIYVGR 252
: || : || : || : || : || : || : || : || : || : || : || : ||
Db 406 ----IVYASLAKGIKTSKVVYFTATFPYV-VIVILLIRGVTLPAGAGIWIYFIPKWEK 460

QY 252 LTRLSIINERFKTQLY 268
|| || || || || || || || || || || || || || || || || || || || ||
Db 461 LTRATVWKDA--ATQIF 475

RESULT 10
US-09-252-991A-16958
; Sequence 16958, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16958
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16958

Query Match          5.6%; Score 80.5; DB 4; Length 150;
Best Local Similarity 40.0%; Pred. No. 1.8;
Matches 20; Conservative 7; Mismatches 22; Indels 1; Gaps 1;

QY 181 YSKVAVALSFLGLWGLGADRFYLCYPALGLKFCUTV:FCGIGSLIDFIL: 236
: || : || : || : || : || : || : || : || : || : || : || : ||
Db 24 HSAKIGLLWIF-GFTGSHRFY:GKPIGTGTFGLFIFGIWIDFLI 72

RESULT 11
US-08-414-657D-44
; Sequence 44, Application US/08414657D
; Patent No. 5861283
; GENERAL INFORMATION:
; APPLICANT: Levitt, Pat
; APPLICANT: Pimenta, Aurea
; APPLICANT: Fischer, Itzhak
; APPLICANT: Zhukareva, Victoria
; TITLE OF INVENTION: Limbic System-Associated Membrane
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dechert Price & Rhoads
; STREET: 997 Lenox Drive, Building 3, Suite 210
; CITY: Lawrenceville
; STATE: NJ
; COUNTRY: USA
; ZIP: 08543
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/414,657D
; FILING DATE: 31-MAR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bloom, Allen
; REGISTRATION NUMBER: 29,135
; REFERENCE/DOCKET NUMBER: 317743-102
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 609-520-3214
; TELEFAX: 609-520-3259
; TELEX:
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 304 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-414-657D-44

Query Match          5.6%; Score 80.5; DB 2; Length 304;
Best Local Similarity 29.9%; Pred. No. 4.8;
Matches 32; Conservative 14; Mismatches 40; Indels 21; Gaps 6;

QY 201 SAGGESLACEDLKVG-----QYICKDPKINDATQEPVNCNTYTAHVSCFPAPNITCKDSS 1:6
: || : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 202 TTGRQASLKCEASAVPADPFWYRDDTRINSANGLEIKSTE---QSSSLVTNVT-EEHY 2:7
|| || || || || || || || || || || || || || || || || || || || ||
QY 157 GNETHTFTGNEVG-----FFKPISCRNVNGYSYKVAVALSFLGLWGL 196
|| || || || || || || || || || || || || || || || || || || || ||
Db 258 GNYTCVANKLGVNANSLVLFHRCGSVRGNG-SISIAVPL-----WL 298
|| || || || || || || || || || || || || || || || || || || || ||

RESULT 12
US-08-414-657D-2
; Sequence 2, Application US/08414657D
; Patent No. 5861283
; GENERAL INFORMATION:
; APPLICANT: Levitt, Pat
; APPLICANT: Pimenta, Aurea
; APPLICANT: Fischer, Itzhak
; APPLICANT: Zhukareva, Victoria
; TITLE OF INVENTION: Limbic System-Associated Membrane
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dechert Price & Rhoads
; STREET: 997 Lenox Drive, Building 3, Suite 210
; CITY: Lawrenceville
; STATE: NJ
; COUNTRY: USA
; ZIP: 08543
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/414,657D
```

;; FILING DATE: 31-MAR-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Bloom, Allen
;; REGISTRATION NUMBER: 29,135
;; REFERENCE/DOCKET NUMBER: 317743-102
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 609-520-3214
;; TELEFAX: 609-520-3259
;; TELEX:

INFORMATION FOR SEQ ID NO: 2:

;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 325 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; FRAGMENT TYPE: internal
US-08-414-657D-2

Query Match 5.6%; Score 80.5; DB 2; Length 325;

Best Local Similarity 29.9%; Pred. No. 5.3;

Matches 32; Conservative 14; Mismatches 40; Indels 21; Gaps 6;

Qy 101 SAGGEESLKCEDLVG-----QYICKDPKINDATQEPVNCNTNYTAHVSCFPAPNITCKDSS 156
Db 223 TTGRQASLKCEASAVPAPDFEYRDDTRINSANGLEIKSTE---GOSSLVTNVT-EEHY 273
Qy 157 GNEHTFTONEVG-----FFKPISCRNVNGYSYKVAVALSLFLGWL 196
Db 279 GNYTCVAANKLGVTNASLVLPFGSVRGING-SISLAVPL-----WL 319

RESULT 13

US-08-414-657D-41
;; Sequence 41, Application US/08414657D
;; Patent No. 5861283
;; GENERAL INFORMATION:
;; APPLICANT: Levitt, Pat
;; APPLICANT: Pimenta, Aurea
;; APPLICANT: Fischer, Itzhak
;; APPLICANT: Zhukareva, Victoria
;; TITLE OF INVENTION: Limbic System-Associated Membrane
;; NUMBER OF SEQUENCES: 60
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Dechert Price & Rhoads
;; STREET: 997 Lenox Drive, Building 3, Suite 210
;; CITY: Lawrenceville
;; STATE: NJ
;; COUNTRY: USA
;; ZIP: 08543

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/414,657D
;; FILING DATE: 31-MAR-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Bloom, Allen
;; REGISTRATION NUMBER: 29,135
;; REFERENCE/DOCKET NUMBER: 317743-102
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 609-520-3214

;; TELEFAX: 609-520-3259
;; TELEX:

INFORMATION FOR SEQ ID NO: 41:

;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 325 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-414-657D-41

Query Match 5.6%; Score 80.5; DB 2; Length 325;

Best Local Similarity 29.9%; Pred. No. 5.3;

Matches 32; Conservative 14; Mismatches 40; Indels 21; Gaps 6;

Qy 101 SAGGEESLKCEDLVG-----QYICKDPKINDATQEPVNCNTNYTAHVSCFPAPNITCKDSS 156
Db 223 TTGRQASLKCEASAVPAPDFEYRDDTRINSANGLEIKSTE---GOSSLVTNVT-EEHY 278
Qy 157 GNEHTFTONEVG-----FFKPISCRNVNGYSYKVAVALSLFLGWL 196
Db 279 GNYTCVAANKLGVTNASLVLPFGSVRGING-SISLAVPL-----WL 319

RESULT 14

US-09-135-080-2
;; Sequence 2, Application US/09135080
;; Patent No. 6423827
;; GENERAL INFORMATION:

;; APPLICANT: Levitt, Pat R.
;; APPLICANT: Pimenta, Aurea
;; APPLICANT: Fischer, Itzhak
;; APPLICANT: Zhukareva, Victoria
;; TITLE OF INVENTION: Limbic System-Associated Membrane
;; NUMBER OF SEQUENCES: 29
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Dechert Price & Rhoads
;; STREET: 997 Lenox Drive, Building 3, Suite 210
;; CITY: Lawrenceville
;; STATE: NJ
;; COUNTRY: USA
;; ZIP: 08543

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/135,080
;; FILING DATE: 17-AUG-1998
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/414,657
;; FILING DATE: 31-MAR-1995

;; ATTORNEY/AGENT INFORMATION:
;; NAME: Bloom, Allen
;; REGISTRATION NUMBER: 29,135
;; REFERENCE/DOCKET NUMBER: 317743-102A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 609-620-3214
;; TELEFAX: 609-620-3259
;; TELEX:

INFORMATION FOR SEQ ID NO: 2:

;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 325 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; FRAGMENT TYPE: internal
US-09-135-080-2

Query Match 5.6%; Score 80.5; DB 4; Length 325;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM protein - protein search, using sw mode:

Run on: September 26, 2003, 17:23:47 ; Search time 27 Seconds
(without alignments)
1507.467 Million cell updates/sec

Title: US-09-852-100A-2

Perfect score: 1439

Sequence: 1 MHILGSPNVIPRAHQKNT.....LHRLSITNE:FRKTQLYF 269

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*

2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*

3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*

4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*

5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*

6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*

7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*

8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*

9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*

10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*

11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*

12: /cgn2_6/ptodata/2/pubpaa/US09A_NEW_PUB.pep.*

13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*

14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*

15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*

16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*

17: /cgn2_6/ptodata/2/pubpaa/US50_NEW_PUB.pep.*

18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1439	100.0	269	9	US-09-852-100A-2
2	1439	100.0	269	10	US-09-833-503A-2
3	1439	100.0	269	15	US-10-199-881-4
4	201	14.0	221	11	US-09-974-879-230
5	201	14.0	222	11	US-09-305-736-230
6	200	13.9	221	10	US-09-833-503A-6
7	200	13.9	221	11	US-09-992-600A-82
8	200	13.9	221	11	US-09-924-340-82
9	200	13.9	221	12	US-09-992-095B-82
10	200	13.9	221	12	US-10-154-678-82
11	200	13.9	221	12	US-09-999-570-82
12	200	13.9	221	15	US-10-000-489-82
13	200	13.9	221	15	US-10-000-986-82
14	200	13.9	221	15	US-10-199-881-6
15	200	13.9	247	11	US-09-796-753-48

16	196	13.6	230	11	US-09-796-753-56	Sequence 50, Appl
17	136.5	9.5	214	10	US-09-833-503A-4	Sequence 4, Appl
18	136.5	9.5	214	15	US-10-199-881-4	Sequence 4, Appl
19	136.5	9.5	224	9	US-09-925-301-1018	Sequence 1018, Af
20	121	8.4	172	11	US-09-798-889-58	Sequence 58, Appl
21	115.5	8.0	213	11	US-09-866-050A-631	Sequence 631, App
22	94.5	6.6	166	15	US-10-156-761-10938	Sequence 10938, A
23	92.5	6.4	87	15	US-10-156-761-10936	Sequence 10936, A
24	90	6.3	443	15	US-10-156-761-9371	Sequence 9371, Ap
25	87	6.0	476	15	US-10-156-761-8693	Sequence 8693, Ap
26	85	5.9	338	10	US-09-808-632-72	Sequence 72, Appl
27	85	5.9	338	11	US-09-800-198-61	Sequence 61, Appl
28	84.5	5.9	389	9	US-09-925-297-489	Sequence 489, App
29	83.5	5.8	179	15	US-10-156-761-13418	Sequence 13418, A
30	83	5.8	764	15	US-10-177-293-27	Sequence 27, Appl
31	82.5	5.7	338	10	US-09-808-602-69	Sequence 69, Appl
32	82.5	5.7	338	11	US-09-800-198-58	Sequence 58, Appl
33	82.5	5.7	338	15	US-10-161-572-58	Sequence 58, Appl
34	82	5.7	416	15	US-10-156-761-10376	Sequence 10376, A
35	81	5.6	717	15	US-10-034-044A-2	Sequence 2, Appl
36	81	5.6	797	9	US-09-795-232-2	Sequence 2, Appl
37	80.5	5.6	373	11	US-09-284-320-17	Sequence 17, Appl
38	80	5.6	328	11	US-09-961-403-2	Sequence 2, Appl
39	80	5.6	328	14	US-10-011-859-28	Sequence 28, Appl
40	80	5.6	476	15	US-10-080-917-7	Sequence 7, Appl
41	80	5.6	660	15	US-10-294-055-8	Sequence 8, Appl
42	79.5	5.5	987	15	US-10-156-761-4515	Sequence 14515, A
43	79	5.5	797	12	US-10-151-813-124	Sequence 124, App
44	78.5	5.5	245	11	US-09-800-198-92	Sequence 92, Appl
45	78	5.4	304	10	US-09-738-626-6200	Sequence 6200, Ap

ALIGNMENTS

Result 1

US-09-852-100A-2

; Sequence 2, Application US/09852100A

; Patent No. US20020058267A1

; GENERAL INFORMATION:

; APPLICANT: American Home Products

; TITLE OF INVENTION: Beta-amyloid Peptide-Binding Proteins and Polynucleotides En

; FILE REFERENCE: AHP981261P2

; CURRENT APPLICATION NUMBER: US/09/852,100A

; CURRENT FILING DATE: 2001-05-09

; PRIOR APPLICATION NUMBER: US 09/172,990

; PRIOR FILING DATE: 1998-10-14

; PRIOR APPLICATION NUMBER: US 60/104,104

; PRIOR FILING DATE: 1998-10-13

; PRIOR APPLICATION NUMBER: PTC/US99/21621

; PRIOR FILING DATE: 1999-10-13

; PRIOR APPLICATION NUMBER: US 09/060,609

; PRIOR FILING DATE: 1998-04-15

; PRIOR APPLICATION NUMBER: US 60/064,583

; PRIOR FILING DATE: 1997-04-16

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2

; LENGTH: 269

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-852-100A-2

Query Match 100.0%; Score 1439; DB 9; Length 269;

Best Local Similarity 100.0%; Pred. No. 2e-135;

Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHILGSPNVIPRAHQKNTRDGTGLYPMRGPFKNLALLPFLGGGSGGSEKYSV 60

|||||#####|

Db 1 MHILGSPNVIPRAHQKNTRDGTGLYPMRGPFKNLALLPFLGGGSGGSEKYSV 60

|||||#####|

QY 61 SKMAAWPSPGSAPEAVTARLVGVFVSVITTPMGAVATSAGGESLKCEDLKVGYIC 120

```
Db 61 SKMAAAMPSPAPAVTARLVGVLFVSVITGPGAVATSGAGEESLKCEDLVGVQYIC 120
QY 121 KDPKINDATQEPVNCNTNTAHVSCFPAPNITCKDSSGNETHFTGNEVGFPPISCRNVNG 180
Db 121 KDPKINDATQEPVNCNTNTAHVSCFPAPNITCKDSSGNETHFTGNEVGFPPISCRNVNG 180
QY 181 YSKYKAVALSFLGWLGDREYLVGYPALGLLKFCFVGCIGSLIDFLLISMQIVGPSDG 240
Db 181 YSKYKAVALSFLGWLGDREYLVGYPALGLLKFCFVGCIGSLIDFLLISMQIVGPSDG 240
QY 241 SSIIDYGTGTRLTSLITNETFRKTQLYP 269
Db 241 SSIIDYGTGTRLTSLITNETFRKTQLYP 269

RESULT 2
US-09-833-503A-2
: Sequence 2, Application US/09833503A
: Patent No. US20020146760A1
: GENERAL INFORMATION:
: APPLICANT: Ozenberger, Bradley A
: APPLICANT: Kajkowski, Eileen M
: APPLICANT: Lo, Chang-Hsiung F
: APPLICANT: American Home Products Corporation
: TITLE OF INVENTION: Polynucleotides Encoded By Them, and Methods of Using
: TITLE OF INVENTION: Polynucleotides Encoded By Them, and Methods of Using
: TITLE OF INVENTION: Same
: FILE REFERENCE: AHP98165-00PCT
: CURRENT APPLICATION NUMBER: US/09/833,503A
: CURRENT FILING DATE: 2000-10-13
: PRIOR APPLICATION NUMBER: 50/104,104
: PRIOR FILING DATE: 1998-10-13
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 2
: LENGTH: 269
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-833-503A-2

Query Match 100.0%; Score 1439; DB 10; Length 269;
Best Local Similarity 100.0%; Pred. No. 2e-135;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHILKSPNVIPIRAHQKNTRRDGTGLYPMRGPFKNLALPFLSLGCGGSGSGEKVS 60
Db 1 MHILKSPNVIPIRAHQKNTRRDGTGLYPMRGPFKNLALPFLSLGCGGSGSGEKVS 60
QY 61 SKMAAAMPSPAPAVTARLVGVLFVSVITGPGAVATSGAGEESLKCEDLVGVQYIC 120
Db 61 SKMAAAMPSPAPAVTARLVGVLFVSVITGPGAVATSGAGEESLKCEDLVGVQYIC 120
QY 121 KDPKINDATQEPVNCNTNTAHVSCFPAPNITCKDSSGNETHFTGNEVGFPPISCRNVNG 180
Db 121 KDPKINDATQEPVNCNTNTAHVSCFPAPNITCKDSSGNETHFTGNEVGFPPISCRNVNG 180
QY 181 YSKYKAVALSFLGWLGDREYLVGYPALGLLKFCFVGCIGSLIDFLLISMQIVGPSDG 240
Db 181 YSKYKAVALSFLGWLGDREYLVGYPALGLLKFCFVGCIGSLIDFLLISMQIVGPSDG 240
QY 241 SSIIDYGTGTRLTSLITNETFRKTQLYP 269
Db 241 SSIIDYGTGTRLTSLITNETFRKTQLYP 269

RESULT 3
US-10-199-881-2
: Sequence 2, Application US/10199881
: Publication No. US20030096356A1
: GENERAL INFORMATION:
: APPLICANT: Wyeth
: TITLE OF INVENTION: No. US20030096356A1 G-Protein-Coupled Receptor-Like Proteins and
```

```
: TITLE OF INVENTION: Encoded by Them, and Methods of Using Same"
: FILE REFERENCE: AHP98165C1
: CURRENT APPLICATION NUMBER: US/10/199,881
: CURRENT FILING DATE: 2002-07-18
: PRIOR APPLICATION NUMBER: PCT/US99/21621
: PRIOR FILING DATE: 1999-10-13
: PRIOR APPLICATION NUMBER: US 90/833,5081
: PRIOR FILING DATE: 2001-12-04
: PRIOR APPLICATION NUMBER: US 60/104,104
: PRIOR FILING DATE: 1998-10-13
: NUMBER OF SEQ ID NOS: 45
: SOFTWARE: Patent In version 3.1
: SEQ ID NO 2
: LENGTH: 269
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-199-881-2

Query Match 100.0%; Score 1439; DB 15; Length 269;
Best Local Similarity 100.0%; Pred. No. 2e-135;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHILKSPNVIPIRAHQKNTRRDGTGLYPMRGPFKNLALPFLSLGCGGSGSGEKVS 60
Db 1 MHILKSPNVIPIRAHQKNTRRDGTGLYPMRGPFKNLALPFLSLGCGGSGSGEKVS 60
QY 61 SKMAAAMPSPAPAVTARLVGVLFVSVITGPGAVATSGAGEESLKCEDLVGVQYIC 120
Db 61 SKMAAAMPSPAPAVTARLVGVLFVSVITGPGAVATSGAGEESLKCEDLVGVQYIC 120
QY 121 KDPKINDATQEPVNCNTNTAHVSCFPAPNITCKDSSGNETHFTGNEVGFPPISCRNVNG 180
Db 121 KDPKINDATQEPVNCNTNTAHVSCFPAPNITCKDSSGNETHFTGNEVGFPPISCRNVNG 180
QY 181 YSKYKAVALSFLGWLGDREYLVGYPALGLLKFCFVGCIGSLIDFLLISMQIVGPSDG 240
Db 181 YSKYKAVALSFLGWLGDREYLVGYPALGLLKFCFVGCIGSLIDFLLISMQIVGPSDG 240
QY 241 SSIIDYGTGTRLTSLITNETFRKTQLYP 269
Db 241 SSIIDYGTGTRLTSLITNETFRKTQLYP 269

RESULT 4
US-09-974-879-230
: Sequence 230, Application US/09974879
: Publication No. US20030028003A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: 125 Human Secreted Proteins
: FILE REFERENCE: P2020P2
: CURRENT APPLICATION NUMBER: US/09/974,879
: CURRENT FILING DATE: 2001-10-12
: PRIOR APPLICATION NUMBER: US 60/239,893
: PRIOR FILING DATE: 2000-10-13
: PRIOR APPLICATION NUMBER: US 09/818,683
: PRIOR FILING DATE: 2001-03-28
: PRIOR APPLICATION NUMBER: US 09/305,736
: PRIOR FILING DATE: 1999-05-05
: PRIOR APPLICATION NUMBER: PCT/US98/23435
: PRIOR FILING DATE: 1998-11-04
: PRIOR APPLICATION NUMBER: US 60/064,911
: PRIOR FILING DATE: 1997-11-07
: PRIOR APPLICATION NUMBER: US 60/064,912
: PRIOR FILING DATE: 1997-11-07
: PRIOR APPLICATION NUMBER: US 60/064,983
: PRIOR FILING DATE: 1997-11-07
: PRIOR APPLICATION NUMBER: US 60/064,900
: PRIOR FILING DATE: 1997-11-07
: PRIOR APPLICATION NUMBER: US 60/064,988
: PRIOR FILING DATE: 1997-11-07
: PRIOR APPLICATION NUMBER: US 60/064,987
: PRIOR FILING DATE: 1997-11-07
```

```

: LOCATION: (184)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-974-879-230

```

166 LSITLGGFGADRFYLGQWEGELGKLFSPGGLGIWTLIDVLLIGVYGVAPADGSLYI 221

[illegible]

[illegible]

Db 138 CTNSTSMTVSCPQRYPANCTVRD----HVHCLGNRT-FPKMLYCNWTTGGYKWSIALA 191
Qy 189 LSLFLGWLGCADREYLGYPALGSLKFCVGFSGIGSLIDFELLSMQLVGFSUSSYI 244
Db 192 LSTITGGFGADREYLGQWRGSLGKLFSGGLGIWTLIDVLEIGVGYVGPADGSLYI 247

Search completed: September 26, 2003, 17:28:52
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:22:11 : Search time 43 seconds
(without alignments)
601.613 Million cell updates/sec

Title: US-09-852-100A-2

Perfect score: 1439

Sequence: 1 MHILKSPNVIPRAHGOKNF.....TLRLTSLTITNEFKTQLYF 263

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	167.5	11.6	573	2 S44605	C02F5.3 protein -
2	159.5	11.1	753	2 T28787	hypothetical prote
3	95	6.6	309	2 H75286	hypothetical prote
4	92	6.4	456	2 S55681	hypothetical prote
5	91	6.3	1324	1 VGIH59	E2 glycoprotein fr
6	89.5	6.2	690	2 T08694	hypothetical prote
7	87.5	6.1	547	2 H75632	Na(+)-linked D-ala
8	87.5	6.1	1207	2 T23754	hypothetical prote
9	85.5	5.9	348	1 S20911	alcohol dehydrogen
10	84	5.8	446	2 T35005	probable integral
11	83.5	5.8	487	2 C70574	probable atrop2 pro
12	83	5.8	137	2 B75447	hypothetical prote
13	83	5.8	348	1 S32521	alcohol dehydrogen
14	83	5.8	764	1 B8H6	component factor
15	83	5.8	1274	2 T10729	transferrin-like p
16	82.5	5.7	338	2 JC4776	limbic-system asso
17	82	5.7	409	2 T03788	laccase (EC 1.10.3
18	81.5	5.7	504	2 T37184	probable monooxyge
19	81.5	5.7	1266	2 AG2695	conserved hypothet
20	81.5	5.7	1266	2 G97477	hypothetical prote
21	81	5.6	443	2 G90388	thermopine protein
22	81	5.6	513	2 A81850	probable integral
23	81	5.6	513	2 C81092	oxalate:formate an
24	81	5.6	758	2 D87369	beta-D-glucosidase
25	81	5.6	1176	2 T49482	hypothetical prote
26	80.5	5.6	134	2 D83595	hypothetical prote
27	80.5	5.6	223	2 F72803	gp12 protein - Myc
28	80.5	5.6	1237	2 A54080	protein-tyrosine-p
29	80	5.6	171	2 T01466	iron-su-fur cofact

ALIGNMENTS

RESULT 1

S44605

C02F5.3 protein - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 02-Feb-2001

C:Accession: S44605

R:Anderson, K.

submitted to the EMBL Data Library, May 1993

A:Description: Sequence of the C. elegans cosmid C02F5.

A:Reference number: S44603

A:Accession: S44605

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-573 <AND>

A:Cross-references: EMBL:L14745; NID:g289607; PID:g289610

C:Genetics:

A:Introns: 224/2; 304/1; 363/3; 390/3; 503/2

C:Superfamily: translation elongation factor Tu homology

C:Keywords: GTP binding; nucleotide binding; P-loop

F:65-183/Domain: translation elongation factor Tu homology <F50>

F:69-76/Region: nucleotide-binding motif A (P-loop)

F:246-249/Region: GTP-binding NKXD motif

Query Match 11.6% Score 167.5; DB 2: Length 573;

Best Local Similarity 27.9%; Pred No. 1.3e-06;

Matches 50; Conservative 23; Mismatches 57; Indels 49; Gaps 5;

QY 90 VTTGPGAVATSAAGSESLKCEDLVQGYICKDP-----K 124

DB 415 VSTNPLGPV-----VECFLENSFLCEDPVPVLYGPGQTGOOPANESFRNEGKCLK 465

QY 125 :NDATQEPVNCNTYAHVSCFPAPNITCKDSNGNETHGTGNEVGFFKPISCRNNVNGYSYK 184

DB 466 MGYRAEDVEFTN-----VACRVLPCTEC---HGPRP-----FTRTPTCIYNGHYEL 510

QY 185 VAVALSLFGLMGLADRFYVLGALFKCTVGCIGSLIDFILSMQIVGSPDGSSY 243

DB 511 TLLYSIFLGVAVDVFCLGYSAMAVGKMLTLGGFGIWMIVDIFLLVLGLGPDADSSM 569

RESULT 2

T28787

hypothetical protein C4D11.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T28787

R:Gattung, S.; Magg1, L.

submitted to the EMBL Data Library, May 1997

A:Description: The sequence of C. elegans cosmid C4D11.

A:Reference number: Z20522

A:Accession: T28787

A>Status: preliminary; translated from GB/EMBL/DDBJ

```
Query Match      6.3%; Score 91; DB 1; Length 1324;  
Best Local Similarity 23.6%; Pred. No. 13;  
Matches 59; Conservative 27; Mismatches 94; Indels 70; Gaps 13:  
  
QY       25 TGLPYPMRG-PFKNLLALP-----FSLPLLGGGSGGEKVSVMARWPSGPAA-- 73  
         ||||| : |:||||| : ||||| : | : | : ||| :  
Db        66 IGYVPDGSKFERNALATGTNSVLSWFQPYPYNDFNGIFAK--VNQLKTSTPGATAYE 123  
  
QY       74 PEAVTARLVGVLMFVSVTTPGMGAVALTSAGSEESLKCEDLKVGOY-ICKDPKNINDAUEP 132  
         ||||| : ||||| : ||| : ||| : ||| :  
Db       124 PTIVIGSLFGYTST-VTWIIPYINGVINAS-----VCQYITCOLP----- 161
```

	Query Match	6.38;	Score 91;	DB 1;	Length 1324;
	Best Local Similarity	23.6%;	Pred. No. 13;		
	Matches	59;	Conservative	27;	Mismatches 94; Indels 70; Gaps 13:
Q7	25	TGLPYPMRG-PFKNNLALLP-----FSLPLLGCGSGGSEKVSVMARAWPSGPQA--	73		
		: : :			:
D6	66	TGYVPDGSKFERNALATGTNSVLSWFQPYPYLNQFNDSIFAQ--VQNLIKSTSPSGATAYE	123		
		: : :			:
QY	74	PEAVTARLVGVLMFVSVTTPGMGAVALTSAGGEESLKCEDLKVGOY-ICKDPKINDATQEP	132		
		: : :			:
D6	124	PTVIGSLFGYTST-VTWIIEPYNGVINMAS-----VCQYITCQLP-----	161		
		: : :			:

QY 133 VNCITYTARYSCFPAPNITCKDSSGNETHFTGNE-VQFF-----KPLSCRNNGYSYKVAV 187
Db 162 -----YTDCKNTN-----GNKLGFHWTDVKKPPICVLKRNFTLVNA 199
QY 188 ALSFLGWLQADRYLYG----PALGLLKFTVGGCGSLIDFLILSMQIVGPSDSSYI 244
Db 200 DAFYFHYQHGTFYAYADKPSATIFLSVY----ICDILTYVYVLPFCINPTAGSTFA 255
QY 245 IDYVGIRLTR 254
Db 256 PRYWTPLVK 265
RESULT 6
T08604
hypothetical protein GRR1 - soybean
C:Species: Glycine max (soybean)
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
C:Accession: T08604
R:Chen, W.; Atherly, A.
submitted to the EMBL Data Library, August 1997
A:Reference number: 215438
A:Accession: T08604
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-690 <CH>
A:Cross-references: EMBL:AF019910; NID:g2407789; PID:g2407790
A:Experimental source: variety L85-3044; root
C:Genetics:
A:Gene: grr1
Query Match
Best Local Similarity 6.2%; Score 89.5; DB 2; Length 690;
Matches 43; Conservative 22; Mismatches 59; Indels 51; Gaps 7;
QY 85 LWFVSITGPMGAVATSAGGSESLKCEDEKYGQYICKDPKINDATQEPV--NCTNYTAHV 142
Db 233 LWDVA-TGVDVGLTEIASGCHOLEKLD-----LCKCPNISDKTLIAVAKNCFN-LAE 283
QY 143 SCFPAPNITCKDSSGNETHFTGNEVGFKPCISCRNVNGYSYKVAVALSLFLCWLQADRFY 202
Db 284 SIETCPNI-----GNEQLQAIKCPNLRISIKNCSGVQ----- 319
QY 203 LGYPALGLLKFTCVGFCGIGSLIDFIL--ISMQIVGPSDSSYI:DNVGTPLTR: 255
Db 320 -----GVAGLLSSASFALTKVKLESITVSDLSAVZGHYGVAVTCL 360

RESULT 7
H75632
Na(+)-linked D-alanine glycine permease - Deinococcus radiodurans (strain M1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: H75632
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, L.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 285, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567255
A:Accession: H75632
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-547 <WH>
A:Cross-references: GB:AF001826; NID:96460827; PIDN:AAH12563.1; PID:96460859; TIGR:DRB01
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRB0133
A:Map position: megapiasmid
A:Genome: plasmid
A>Note: plasmid MPI
C:Superfamily: sodium-dependent D-alanine/glycine transport protein.

Query Match
Best Local Similarity 6.1%; Score 87.5; DB 2; Length 547;
Matches 57; Conservative 19; Mismatches 52; Indels 95; Gaps 13;
QY 43 SLPLLGGGGSGSGEKYSVSKMAA-----WPS--GFSAPAEVZARLVGV----- 84
Db 29 SRPLSSESSSSAQEPWGRPLPAALVFTGLLGAVSASQAQSPVDERINAVVTPVSHFLS 88
QY 85 -LWFVSVTGTP-----WGAVATSAGGSESLKCE-----DLKVGQYIC 120
Db 89 GLIFASISVGEAQVPLIVVWLAVA-----AVCTLSLRFVNIWGFKHGIDLVGRGY-- 139
QY 121 KDPKINDATQEPVNCNTNYTAHVSCEPAPNITCKDSSGNETHFTGNEVGFKPCISCRNVNG 180
Db 140 ----GNDA-----DAGWTHFVALTAVSSTVGLGNIG 170
QY 181 YSYKVAVALSLFLCWLQADRFYLYGYPALGLL----KF--CTVG 217
Db 171 ----VAVALS--LGGPGATFMMI---LVGLLSMSTKFEVCTLG 204
RESULT 8
T23754
hypothetical protein T05C12.10 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T23754; T24513
R:Thomas, K.
submitted to the EMBL Data Library, June 1995
A:Reference number: Z19793
A:Accession: T23754
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1207 <MIL>
A:Cross-references: EMBL:Z49968; PIDN:CAA90265.1; GSPDB:GN00020; CESP:T05C12.10
A:Experimental source: clone M110
R:Burton, J.
submitted to the EMBL Data Library, October 1995
A:Reference number: Z19901
A:Accession: T24513
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1207 <M12>
A:Cross-references: EMBL:Z66500; PIDN:CAA91313.1; GSPDB:GN00020; CESP:T05C12.10
A:Experimental source: clone T05C12
C:Genetics:
A:Gene: CESP:T05C12.10
A:Map position: 2
A:Mutrons: 31/3; 87/2; 141/3; 160/2; 203/3; 267/1; 776/2; 794/2; 834/2; 1086/3; 1143
Query Match
Best Local Similarity 6.1%; Score 87.5; DB 2; Length 1207;
Matches 47; Conservative 26; Mismatches 83; Indels 67; Gaps 8;
QY 2 HILKSPNPVIPRAHGOKNRDGTGLYPMRGPKFNALLPFLPLGGGGSGSGEKYSVS 61
Db 848 HNAESSASGIPLYQARSNTYNGAPVPAPGS-----GATSGTSGSGTSESVT 896
QY 62 KMAANWPSGFS-----APEAVTARLVGV:LWFVSVTGPMGAVATSAG 100
Db 897 NGSGATESGSGTSGTGTGTSGTSGSGTGSAAARTSSIAGDAPQAVALADTPGAAGAAG 956
QY 104 GE-----ESL-----KCEDLKVGQYICKDPKINDAT-----QEPVNCNTNYTA 140
Db 957 GRSNCFSSADSLVTITVGQRMDLQIGDYLVLPSSGNVLKYKVEFYHREPKTRNF-- 1054
QY 141 HVSCFPAPNITCKDSSGNETHFTGNEVGFKPCIS--CRNVNGYS 182
Db 1025 -----VLYTKSGRKLTLGRHL---LPVAECSSQVEQYT 1045
RESULT 9
S20911

```

CY      39  LIPFSLPLGGGSGGKVSYSKMAAAMPSSGSAPEAVIAHLVGVLWF--VSVTGPWG 96
       1 : 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Cb      106  LVLVGGFVGIGLIGYISPVSSTLIKWPDRPG-----MATGIAIMGGGALIASPWS 160
       1 : 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
CY      97  AVATSAGGEE-----LKCDLKVQYICKDPK-----IN 126
       1 : 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Cb      161  KQMLKSGFDNSGIALAFVLVHGLTAYVFMGLLVLYVRPRERADGRPALEGVQVSAR 220
       1 : 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
CY      127  DATQEP-----VNCNTYIAHVSCEF--PAPNTIKDSSCNETHFTGNEGVFFKPISCRN 177
       1 : 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Cb      221  SAVRTQFLLMIVLCMNVYTAGIGILEKAAPMTDFSDTSPVSYTAAAGFVALLSAAN 280
       1 : 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
CY      178  VNGYSYKVAVALSLFLGW-----LCADRFYLGYPALGILLKCTGVGFCGIGSLIPILIS 231
       1 : 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Cb      281  MAGR-----FGWSASDILGRKNLYRYLVGVALMYTLTALFGDSSKPLFVLC 329
       1 : 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
CY      232  MOIV-----GPSDGSYYIIDYGT 250
       1 : 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Cb      330  LVVSPYGGGFATAPAYLKDLFGT 353
       1 : 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11

RESULT 11
C70574
C:Probable arop2 protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: C70574
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, J.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Harrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome.
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: C70574
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-487 <COL>
A:Cross-references: GB:295324; GB:AL123456; NID:g2261760; PIDN:CA808578.1; PID g20542
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: arop2
C:Superfamily: arginine permease

Query Match 5.8%; Score 83.5; DB 2; Length 487;
Best Local Similarity 26.5%; Pred. No. 19;
Matches 27; Conservative 16; Mismatches 40; Indels 19; Gaps 5.

CY      159  ETHFTGNEGVFFKPISCRNVNGYSYKVAVALSLFLGWLCAADRFLYLGYPALGILLKCTGVG 218
       1 : 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Cb      8  DEPLTREDYGHKGLHSRQLMIALGCGATGTLFG--AGGRLASAGPSG----FLVYGI 61
       1 : 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
CY      219  CGIGSLIDFILISMVQV-----PSDGS--SYIIDYGTSL 252
       1 : 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Cb      62  CGI-----EVFLTLRALGELVLHRPSSGGSFVSAREFYGEKV 98
       1 : 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11

RESULT 12
B75447
C:Hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: B75447
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: B75447
A:Status: preliminary
A:Molecule type: DNA

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:15:17 : Search time 23 Seconds
(without alignments)

550.008 Million cell updates/sec

Title: US-09-852-100A-2

Perfect score: 1439

Sequence: 1 MHILKSPNVPVRAHGKQNT.....IRLRLSLTNEIRKTKQLYP 259

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	167.5	11.6	573	1 YKK3_CAEEL	P34280 caenorhabdi
2	90	6.3	1324	1 VGL2_CVNA5	P11224 murine coro
3	89.5	6.2	515	1 EF15_PORPU	P50257 porphyra pu
4	86.5	6.0	151	1 LCT2_MOUSE	O88803 mus musculu
5	85.5	5.9	348	1 ADH2_KLUJA	P49383 kluyveromyc
6	83.5	5.9	338	1 LAMP_RAT	Q62813 rattus norv
7	83.5	5.8	487	1 Y346_MYCTU	O06297 mycobacteri
8	83	5.8	349	1 ADH1_KLUJA	Q67289 kluyveromyc
9	82.5	5.7	338	1 CFAB_HUMAN	P00751 homo sapien
10	82.5	5.7	338	1 LAMP_HUMAN	O13445 homo sapien
11	80.5	5.6	223	1 VG32_BPMO2	O84226 mycobacteri
12	80.5	5.6	455	1 EN11_HUMAN	Q99805 homo sapien
13	80	5.6	328	1 IBP2_HUMAN	P18065 homo sapien
14	80	5.6	450	1 LIPP_PIG	P00591 sus scrofa
15	80	5.6	638	1 OARI_LYMAST	O77408 lymanaca sta
16	79.5	5.5	489	1 ANSP_MYCTU	Q33261 mycobacteri
17	79.5	5.5	655	1 POLG_HRV3	Q82081 human rhino
18	79.5	5.5	910	1 AUX1_BOVIN	Q27974 bos taurus
19	79.5	5.5	3432	1 POLG_JAEV1	P27395 j domestico po
20	79.5	5.5	3432	1 POLG_JAEV5	P19110 j genome po
21	79.5	5.5	3432	1 POLG_JAEV7	P23886 j genome po
22	79.5	5.5	4367	1 DYHC_NEUCR	P45443 neuropept
23	79	5.5	468	1 LIP2_MOUSE	P17892 mus musculu
24	79	5.5	525	1 STP_SPIOL	Q03411 spinacia ol
25	79	5.5	797	1 SGA5_HUMAN	Q9Y345 homo sapien
26	78.5	5.5	245	1 CR53_HORSE	O19010 equus caball
27	78.5	5.5	398	1 MUB1_XENLA	P38565 xenopus lae
28	78.5	5.5	5376	1 ZAN_MOUSE	O88799 mus musculu
29	78	5.4	557	1 YP85_MYCTU	Q50636 mycobacteri
30	77.5	5.4	374	1 AKOC_METTH	O26843 methanobact
31	77.5	5.4	2179	1 POLG_HRV14	P03303 human rhino
32	77	5.4	365	1 IDI2_METMA	Q6pw37 methanosarc
33	77	5.4	1302	1 NRG_DROME	P20241 drosophila

34	76.5	5.3	772	1 PMIP_YEAST	P35999 saccharomyc
35	76	5.3	379	1 METX_NEIMB	O9J295 neisseria m
36	76	5.3	467	1 LIPI1_HUMAN	P54315 homo sapien
37	76	5.3	484	1 GATA_PSEAE	Q9hvt8 pseudomonas
38	76	5.3	796	1 STR3_MOUSE	C9erg2 mus musculu
39	76	5.3	1556	1 VG67_HSV1	O00107 ictaluriid h
40	75.5	5.2	199	1 WRBA_YERPE	Q82f61 yersinia pe
41	75.5	5.2	21	1 CLD7_MOUSE	Q92261 mus musculu
42	75.5	5.2	350	1 ADH1_KLUJA	P20369 kluyveromyc
43	75.5	5.2	484	1 PER2_VOLCA	P81132 volvox cart
44	75.5	5.2	592	1 CEA_CITFR	P04480 citrobacter
45	75.5	5.2	656	1 PER_DROSI	Q03355 drosophila

ALIGNMENTS

RESULT 1
YKK3_CAEEL
ID YKK3_CAEEL STANDARD: PKT; 573 AA.
AC P34280; 573 AA.
CT 01-FEB-1994 (Rel. 28, Created)
CT 01-FEB-1994 (Rel. 28, Last sequence update)
CT 28-FEB 2003 (Rel. 41, Last annotation update)
DE Hypothetical GTP-binding protein C02F5.3 in chromosome III.
GN C02F5.3
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID:6239;
RN [1]
RP SEQUENCE FROM N.A.
NC STRAIN-Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lighthouse J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkhen R.,
RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Weidman P.;
R "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
R elegans";
RL Nature 368:32-38(1994).
CC -!- S-MILARITY: Belongs to the GTP1 / OBG family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L14745; AAA27918.1;
CC PIR; S44605; S44605.
CC WormPep; C02F5.3; CE00039.
CC InterPro; IPR006074; GTP1/OBG_dom.
CC InterPro; IPR006073; GTP1_OBG.
CC InterPro; IPR006169; GTP1_OBG_sub.
CC InterPro; IPR005225; Small_GTP.
CC Pfam; PF01018; GTP1_OBG; 1.
CC Pfam; PF02824; TGS; 1.
CC Pfam; PF05154; TM2; 1.
CC PRINTS; PR00326; GTP1_OBG.
CC TIGRFAMs; TIGR00231; small_gtp; 1.
CC PROSITE; PS00905; GTP1_OBG; 1.
KW Hypothetical protein; GTP-binding.

FT	CARBOHYD	435	435	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	530	530	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	625	625	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	657	657	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	665	665	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	698	698	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	737	737	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	754	754	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	893	893	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	1180	1180	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	1190	1190	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	1209	1209	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	1225	1225	N-LINKED (GLCNAC..)	(POTENTIAL).
FT	CARBOHYD	1246	1246	N-LINKED (GLCNAC..)	(POTENTIAL).
SEQ	SEQUENCE	1324 AA;	146019 MW;	EL96EF8F0BCDBF0E	CRC64;

Query Match: 6.3%; Score 90; DB 1; Length 1324;
Best Local Similarity 23.6%; Pred. No. 5.6;
Matches 59; Conservative 27; Mismatches 94; Indels 70; Gaps 13

QY	25	TGLYPMRG-PFKNALLP-----	PSLPLGGGGSGGEKYSVSKMAAWPSGSA--	73
DB	66	TGYYPVDSGKFNALRGTNSVSWFPPYLNQDIFAK--	VQMLKTIPTSGATAYF	123
QY	74	PEAVTARVLGVLFVSVTGPWGAVATSSAGEESLKGEDLVKGQY-ICKQPKINDATQEP	132	
DB	124	PRIVIGSLFGYTSY-TVVIEPYNGVMAAS-----	VCQYTTICLP-----	161
QY	133	VNCTNYTAHVSCFPAPNITCKDSSGNETHFTGNE-VGFF-----	KPISCRNVNGYSYKVAV	197
DB	162	-----YTDCCKPNTN-----	GNKLIGFWHTDVKPP:CVLKRNFITLVNA	199
QY	188	ALSFLGMLGADRFVLGY---PALGLLKFCVTGFCGICSLIDFLISMQVIGPSDGSYYI	244	
DB	200	DAFYHFHYOHGGTFYAYYADKPESATITELFSVY-----	IGDILTQYVYLPFCINPTAGSIPA	255
QY	245	IDYVGTRLTR	254	
DB	256	PRYVWTPLVK	265	

RESULT 3
EFIS_PORPU STANDARD: PRT; 515 AA.
AC P50257;
DT 01-OCT-1996 (Rel. 34, Created)
CT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
FE Elongation factor :alpha S (EF-1-alpha S) (Sporophyte-specific EF-1-alpha)
DE alpha)
GN tEF-S.
OS Porphyra purpurea.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Porphyra.
QX NCBI_TaxID:2787;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Avonport;
RX MEDLINE-96309385; PubMed-8704161;
RA Liu Q.Y., Baldauf S.L., Reith M.E.;
RT "Elongation factor 1 alpha genes of the red alga Porphyra purpurea include a novel, developmentally specialized variant.";
RT Plant Mol. Biol. 31:77-85(1996).
RL
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF AMINOACYL-tRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED ONLY IN THE SPOROPHYTE, A SHELL-BORING, FILAMENTOUS PHASE.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY. EF-TU/EF-1A SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC EMBL: U08841; AAA61750.1; -.
CC HSSP: P07157; IALP.
CC InterPro: IPR004539; EFL_alpha.
CC InterPro: IPR000795; EFU_Grpbid.
CC InterPro: IPR004160; EFU_Cterm.
CC InterPro: IPR004161; EFU_D2.
CC Pfam: PF00009; GTP_EFTU; 1.
CC Pfam: PF03144; GTP_EFTU_D2; 1.
CC Pfam: PF03143; GTP_EFTU_D3; 1.
CC TIGRFAMs: TIGR00483; EF-1_alpha; 1.
CC PROSITE: PS00301; EFATOR_GTP; 1.
CC Elongation factor; protein biosynthesis; GTP-binding;
CC Multigene family.
FT NP_BIND 14 21 GTP (BY SIMILARITY).
FT NP_BIND 91 95 GTP (BY SIMILARITY).
FT NP_BIND 151 154 GTP (BY SIMILARITY).
FT NP_BIND 154 154 GTP (BY SIMILARITY).
SQ SEQUENCE 515 AA; 56648 MW; ERA3F4029F62350 CRC64;

Query Match 6.2%; Score 89.5; DB 1; Length 515;
Best Local Similarity 24.7%; Pred. No. 2.1;
Matches 37; Conservative 20; Mismatches 56; Indels 37; Gaps 6;

QY 41 PPSLP-----LGGGSGSGEKVSVKMAAAW-----PSG/SAPFVNTARLVGVWVSV 90
DB 261 PLRLPLQVDYKIGGIGTVPVGRVETGILKAGMQVTEPAGKAAVEKSVEM-----HH 313
QY 91 ITGPGAVATSGAGGESLKCDLKVGVICDKPK-----INDATOPPVNCTN-- 137
DB 314 TSVPQAIQGDVNGVFNKVLIVDKRGG-VCGDTKNDPPVTECFIANVLIQDHKNRNGY 372
QY 138 -----YTAHVSCEFPAPIITCKDDSSGNETH 161
DB 373 TPVLDCHTAHIAKPFASILSKKDKRGKQTH 402

RESULT 4
LCT2_MOUSE
ID LCT2_MOUSE STANDARD; PRT: 151 AA.
AC O88803; O88804; O90WN3; Q9Z337;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Leukocyte cell-derived chemotaxin 2 precursor (Chondromodulin II)
DE (Chn-II).
GN LECT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANT VAL-123.
RC STRAIN=BAJ/c; TISSUE=Liver;
RX MEDLINE=98382586; PubMed=9714793;
RA Yamagoe S., Watanabe I., Mizuno S., Suzuki K.;
RT "The mouse lect2 gene: cloning of cDNA and genomic DNA, structural
characterization and chromosomal localization.";
RL Gene 216:171-178(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=Swiss Webster / NIH; TISSUE=Embryo, and liver;
RX MEDLINE=99160594; PubMed=1005029;
RA Shukunami C., Kondo J., Wakai H., Takahashi K., Inoue H., Kamizono A.,
RA Hiraki Y.;
RT "Molecular cloning of mouse and bovine chondromodulin-II cDNAs and the
growth-promoting actions of bovine recombinant protein.";
RL J. Biochem. 125:436-442(1999).
CC -!- FUNCTION: Has a neutrophil chemotactic activity. Also a positive
regulator of chondrocyte proliferation.
-!- SUBCELLULAR LOCATION: Secreted.
-!- ALTERNATIVE PRODUCTS:
Event-Alternative splicing: Named isoforms=2;
Name=1; Synonyms=LECT2;
Name=2; Synonyms=LECT2Q;
IsoId=O88803-1; Sequence=Displayed;
IsoId=O88803-2; Sequence=VSP_003051;
-!- TISSUE SPECIFICITY: Highly expressed in liver and weakly in
testis. Not expressed in heart, brain, spleen, lung, skeletal
muscle and kidney.
-!- SIMILARITY: BELONGS TO THE LECT2 / MIM-1 FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation,
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```

```
CC -|- COFACTOR: Binds 2 zinc ions per subunit (By similarity).
CC -|- SUBUNIT: Homotetramer (By similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- SIMILARITY: Belongs to the zinc-containing alcohol dehydrogenase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X64397; CA445739.1; -.
CC FIRM: S20911; S20911.
CC InterPro: IPR002328; ADH_zinc.
CC InterPro: IPR002085; ADH_zn_family.
CC Pfam: PF00107; ADH_zinc_N; 1.
CC PROSITE: PS00059; ADH_ZINC; 1.
CC Oxidoreductase; Zinc; Metal-binding; NAD: Multicene family.
CC METAL 44 44 ZINC 1 (CATALYTIC) (BY SIMILARITY).
CC METAL 67 67 ZINC 1 (CATALYTIC) (BY SIMILARITY).
CC METAL 98 98 ZINC 2 (BY SIMILARITY).
CC METAL 101 101 ZINC 2 (BY SIMILARITY).
CC METAL 104 104 ZINC 2 (BY SIMILARITY).
CC METAL 112 112 ZINC 2 (BY SIMILARITY).
CC METAL 154 154 ZINC 1 (CATALYTIC) (BY SIMILARITY).
CC METAL 348 AA; 37097 MW; F3B64AE1F520689C CRC64;
CC SEQUENCE 348 AA; 37097 MW; 5.9%; Score 85.5; DB 1; Length 348;
Query Match: 5.9%; Score 85.5; DB 1; Length 348;
Best Local Similarity 19.9%; Pred. No. 2.9;
Matches 65; Conservative 35; Mismatches 111; Indels 115; Gaps 13;
QY 19 NTRDQ---TGILPMRGPPKNAALPFSLPLOGGSGSGXSVSKMAAW----- 67
DB 37 NVKYGVCHTDLHAWGDWP---LPTKLPLV-GGHEGAGVYVAMGVKGMIIIGLFA 91
QY 68 -----PSGPSAPEAVT 75
DB 92 KWLNGSCMSECELSNESNCPDADLSGYTHDGSFQYATADAVQAARIPKGTU-AEVA 151
QY 79 ARLVG-----LMFVSVTGPGWAVATSGGSESLKCEDLV----- 115
DB 152 ILCAGVTVVYKALKSADLKAGDVAISGACGGGSGIAIOYAKAMGYRVGLTCAEKALF 211
QY 116 ---GGYICKPKINGATCEPVNCZNYTAH-----VSCFPAFNITCKDSSGNETHTCN 165
DB 212 KELGGYFYDYAVSKDITREIVDATNGAHGV-NVSVSEFAI-----ECSTNYVRSNG 265
QY 166 EVGFFPKPIPCRNNGYSYKVAVALSLFLQWLG--AD-RFYVIGYFALGL-KFCTVGFQGLG 222
DB 266 VLVGLPRDAKCKSDVFTQVVKSVISVGVNGKADTREALDFANGLV-HAPIKIVGLS 324
QY 223 SLIDFI--LISMGIVGPSOSSYIID 246
DB 325 ELADYVYKMYKGEIVG-----RYVVD 345
RESULT 6
LAMP_RAT STANDARD: PRT: 338 AA.
AC Q62813;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Limbic system-associated membrane protein precursor (LSAMP).
GN LSAMP OR LAMP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
```

```

FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 163 183 POTENTIAL.
FT TRANSMEM 214 234 POTENTIAL.
FT TRANSMEM 256 276 POTENTIAL.
FT TRANSMEM 290 310 POTENTIAL.
FT TRANSMEM 341 361 POTENTIAL.
FT TRANSMEM 369 389 POTENTIAL.
FT TRANSMEM 414 434 POTENTIAL.
FT TRANSMEM 440 460 POTENTIAL.
SQ SEQUENCE 487 AA: 52194 MW: 3572502DB6ACD987 CRC64;

Query Match 5.8%; Score 83.5; DB 1; Length 487;
Best Local Similarity 26.5%; Pred. NO. 6.5;
Matches 27; Conservative 16; Mismatches 40; Indels 19; Gaps 5;

QY 159 ETHFTGNEVGFEKPKISCRNVNGYSYKVAVALSLFLGLWGLADREYLCYPAIGLLKFKCTVG 218
DB 8 DERLTREDTGYNKGLHSROLMIALGALGTGFLG--ASGRIASAGPGL---FLVYGI 61
QY 219 CGTGLSLIDFILISMQIVG-----PSDGS--SYIIDYYGTRI 252
DB 62 CGI----FVFLILRALGELVLRHPSGGSFVSAREFYGEKV 98

RESULT 8
ADH1_KLUMA STANDARD: PRT; 348 AA.
TO ADH1_KLUMA
AC Q07388;
CT 01-FEB-1995 (Rel. 31, Created)
CT 01-FEB-1995 (Rel. 31, Last sequence update)
CT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Alcohol dehydrogenase 1 (EC 1.1.1.1).
GN ADH1.
OS Kluyveromyces marxianus (Yeast) (Kluyveromyces fragilis).
OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
NCBI_TaxID=4911;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12424;
RC MEDLINE=93350057; PubMed=8485163;
RA Ladriere J.M., Deicour J., Vandenhaute J.;
RT "Sequence of a gene coding for a cytoplasmic alcohol dehydrogenase
RT from Kluyveromyces marxianus ATCC 12424.";
RL Biochim. Biophys. Acta 1173:99-101(1993).
CC !- CATALYTIC ACTIVITY: An alcohol + NAD(+) -> an aldehyde or ketone +
CC NADH.
CC !- COFACTOR: Binds 2 zinc ions per subunit (By similarity).
CC !- SUBUNIT: Homotetramer.
CC !- SUBCELLULAR LOCATION: Cytoplasmic.
CC !- SIMILARITY: Belongs to the zinc-containing alcohol dehydrogenase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X60224; CAA42785.1; -.
CC PIR: S32521; S32521.
CC InterPro: IPR002328; ADH_zinc.
CC InterPro: IPR002085; Adh_zn_family.
CC Pfam: PF00107; ADH_zinc.N; 1.
CC PROSITE: PS00059; ADH_ZINC; 1.
CC Oxidoreductase; Zinc; Metal-binding; NAD; Multigene family.
CC METAL 44 44 ZINC 1 (CATALYTIC) (BY SIMILARITY).
CC METAL 67 67 ZINC 1 (CATALYTIC) (BY SIMILARITY).
CC METAL 98 98 ZINC 2 (BY SIMILARITY).
CC METAL 101 101 ZINC 2 (BY SIMILARITY).
CC METAL 104 104 ZINC 2 (BY SIMILARITY).
CC METAL 104 104 ZINC 2 (BY SIMILARITY).

```

```

FT METAL 112 112 ZINC 2 (BY SIMILARITY).
FT METAL 154 154 ZINC 1 (CATALYTIC) (BY SIMILARITY).
SQ SEQUENCE 348 AA: 37158 MW: A75D2EBE82E35BD CRC64;

Query Match 5.8%; Score 83; Dh 1; Length 348;
Best Local Similarity 20.8%; Pred. No. 4.9;
Matches 65; Conservative 38; Mismatches 107; Indels 102; Gaps 14;

QY 19 NTRDQ---TGIVPMRGPKKIALLPSPULPLGGSGSGGEKVSYSKMAAAPSG----- 70
Db 37 NVKYSVCHTDLHAWGDWP-----LDTKPLV-GEHAGAGIVVANGENVTEIGDYGAGI 9;

QY 71 -----PSAPEA-----VTARLV----- 82
Db 92 KWLNGSCMSCECELSNCPKADLSGYTHDGSFOQYATADAVOAAIPKKNVDLAEPV 151

QY 83 ----GV-----LWFVSVTGPMCAVATSAGGEESLKCDLKV----- 115
Db 152 ILCAGVTYVYKALKSAHIKAGDWVATSGAGCGGLSLAIQYAKAMGYRVLGIDAGDEKAKLF 211

QY 116 ----GOYICKDKINDATPEVNCINVTYAHVSCFPAPNITCKDSGNFTFTGNE--VGF 169
Db 212 KELGGEYFDFTKDKMVAEIVATNGVAHVINVSSEAAISTSVLYTRSGTVVLVGL 27;

QY 170 FPIPCRNNGVYKYVAVALSLFLWLG--AD-RFYLVPAIGLLK--FCTVGFQGTGSL 224
Db 272 PRDAOCK--SDFNVQVKSISVGVGNRADTREALDFRSGLVKAP-KILGLSELASV 329

QY 225 IDPILISMIOIVG 236
Db 330 YD-KMKVKGQIVG 340

RESULT 9
CEAB_HUMAN STANDARD; PRT: 764 AA
AC P00751; O15006; Q29944; Q96HX6; Q9BTF5; Q9BX92;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement factor B precursor (EC 3.4.21.47) (C3/C5 convertase)
DE (Properdin factor B) (Glycine-rich beta glycoprotein) (GBG) (P3B2).
GN BF.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ALLELES S; FA AND FB).
RX MEDLINE=91065702; PubMed=2249879;
RA Davrinche C., Abbai M., Clerc A.;
RC "Molecular characterization of human complement factor B subtypes."
RL Immunogenetics 32:309-312(1990).
RN [2]
RP SEQUENCE FROM N.A. (ALLELE S).
RC TISSUE=Liver;
RX MEDLINE=94237735; PubMed=8181962;
RA Mejia J.E., Jahn I., de la Salle H., Hauptmann G.;
RT "Human factor B. Complete cDNA sequence of the RF*S allele."
RL Hum. Immunol. 39:49-53(1994).
RN [3]
RP SEQUENCE FROM N.A. (ALLELE S).
RC TISSUE=Liver;
RX MEDLINE=94041399; PubMed=8225386;
RA Schwaible W., Luettig B., Sokolowski T., Estalier C., Weiss E.H.,
RT Meyer Zum Bueschenfelde K.-H., Whaley K., Dippold W.;
RT "Human complement factor B: functional properties of a recombinant
RL zymogen of the alternative activation pathway convertase."
RN Immunobiology 188:221-232(1993).
RN [4]
RP SEQUENCE FROM N.A. (ALLELE S).
RX MEDLINE=94067177; PubMed=8247029;
RA Horiuchi T., Kim S., Matsumoto M., Watanabe I., Fujita S.,

```

```

RA Volanakis J.E.;
RT "Human complement factor B: cDNA cloning, nucleotide sequencing,
RT phenotypic conversion by site-directed mutagenesis and expression."
RL Mol. Immunol. 30:1587-1592(1993).
RN [5]
RP SEQUENCE FROM N.A.
RA Rwen L., Dankers C., Baskin D., Faust J., Loretz C., Alheart M.E.,
RA Banta A., Swartzell S., Smith T.M., Spies T., Hood L.;
RT "Sequence determination of 300 kilobases of the human class III MHC
RT locus."
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins B.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Aleschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rata S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.T., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [7]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Jaatinen T., Kanerva J., Poutanen K.E., Saarinen-Pihkala U.,
RA Lokki M.-L.;
RT "Expression and alternative splicing of human factor B gene in
RT leukemic mononuclear cells."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE OF 26-764, PARTIAL SEQUENCE FROM N.A., AND CARBOHYDRATES.
RX MEDLINE=84161997; PubMed=6546754;
RA Mole J.E., Anderson J.K., Davison E.A., Woods D.E.;
RT "Complete primary structure for the zymogen of human complement
RT factor B."
RL J. Biol. Chem. 259:3407-3412(1984).
RN [9]
RP SEQUENCE OF 260-764.
RX MEDLINE=83204002; PubMed=6342610;
RA Christie D.L., Gagnon J.;
RT "Amino acid sequence of the Bb fragment from complement Factor B.
RT Sequence of the major cyanogen bromide-cleavage peptide (CB-II) and
RL completion of the sequence of the Bb fragment."
RN Biochem. J. 209:61-70(1983).
RN [10]
RP SEQUENCE OF 339-764 FROM N.A.
RX MEDLINE=83273641; PubMed=6308626;
RA Campbell R.D., Porter R.R.;
RT "Molecular cloning and characterization of the gene coding for human
RL complement protein factor B."
RN Proc. Natl. Acad. Sci. U.S.A. 80:4464-4468(1983).
RN [11]
RP SEQUENCE OF 467-595 AND 752-764 FROM N.A.
RX MEDLINE=83039428; PubMed=6957884;
RA Woods D.E., Markham A.F., Ricker A.T., Goldberger G., Colten H.R.;
RT "Isolation of cDNA clones for the human complement protein factor B,
RT a class III major histocompatibility complex gene product."
RL Proc. Natl. Acad. Sci. U.S.A. 79:5661-5665(1982).
RN [12]
RP SEQUENCE OF 16-259 FROM N.A.

```


RX MEDLINE-84158524; PubMed-6323161;
RA Morley B.J., Campbell, R.D.;
RI "Internal homologies of the Ba fragment from human complement
RI component Factor B, a class III MHC antigen.";
RN EMBO J. 3:153-157(1984).
RN [13]
RP SEQUENCE OF 1-99 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE-87102880; PubMed-3643061;
RA Wu L.-C., Morley B.J., Campbell R.D.;
RI "Cell-specific expression of the human complement protein factor B
RI gene: evidence for the role of two distinct 5'-flanking elements.";
RN Cell 48:331-342(1987).
RN [14]
RP GLYCATION OF LYS-291.
RX MEDLINE-91174758; PubMed-2006911;
RA Nlemann M.A., Shown A.S., Miller E.J.;
RI "The principal site of glycation of human complement factor H.";
RN Biochem. J. 274:473-480(1991).
CC -!- FUNCTION: FACTOR B WHICH IS PART OF THE ALTERNATE PATHWAY OF THE
CC COMPLEMENT SYSTEM IS CLEAVED BY FACTOR D INTO 2 FRAGMENTS: BA AND
CC BB. BA, A SERINE PROTEASE, THEN COMBINES WITH COMPLEMENT FACTOR IS
CC TO GENERATE THE C3 OR C5 CONVERTASE. IT HAS ALSO BEEN IMPLICATED
CC IN PROLIFERATION AND DIFFERENTIATION OF PREACTIVATED B
CC LYMPHOCYTES, RAPID SPREADING OF PERIPHERAL BLOOD MONOCYTES.
CC STIMULATION OF LYMPHOCYTE BLASTOGENESIS AND LYSIS OF ERYTHROCYTES.
CC BA INHIBITS THE PROLIFERATION OF PREACTIVATED B LYMPHOCYTES.
CC -!- CATALYTIC ACTIVITY: Cleaves C3 in the alpha-chain to yield C3a and
CC C3b. Cleaves C5 in the alpha-chain to yield C5a and C5b. Both
CC cleavages take place at the C-terminal of an arginine residue.
CC -!- SUBUNIT: Monomer.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing: Named isoforms-2;
CC Name=1;
CC IsoId=P00751-1; Sequence-Displayed;
CC Name=2;
CC IsoId=P00751-2; Sequence-VSP_005380, VSP_005381;
CC -!- POLYMORPHISM: TWO MAJOR VARIANTS, F AND S, AND 2 MINOR VARIANTS,
CC AS WELL AS AT LEAST 14 VERY RARE VARIANTS, HAVE BEEN IDENTIFIED.
CC -!- MISCELLANEOUS: FACTOR B IS A MAJOR HISTOCOMPATIBILITY COMPLEX
CC CLASS-III PROTEIN.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 3 Sushi (SCR) domains.
CC -!- SIMILARITY: Contains 1 WUFA domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X72875; CAA51389.1; -;
DR EMBL: S67310; AAD13989.1; -;
DR EMBL: L15702; AAA16820.1; -;
DR EMBL: X00284; CAA25077.1; -;
DR EMBL: AF019413; AAB67977.1; -;
DR EMBL: BC004143; AAH04143.1; -;
DR EMBL: BC007990; AAH07990.1; -;
DR EMBL: AF349679; AAK30167.1; -;
DR EMBL: K01566; AAA36225.2; -;
DR EMBL: J00125; -; NOT_ANNOTATED_CDS.
DR EMBL: J00126; AAA36226.1; -;
DR EMBL: J00185; AAA36219.1; -;
DR EMBL: J00186; AAA36220.1; -;
DR EMBL: M15082; AAA59625.1; -;
DR EMBL: PIR; S34075; BBHU.
DR HSSP: P20231; IAAO.
DR MEROPS: S01.196;
DR SWISS-2DPAGE: P00751; HUMAN.
DR Slena-2DPAGE: P00751; -;
DR Genew: HGNC:1037; BF.

DR MIM: 138470; -;
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser. protease. Try.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF00084; sushi; 3.
DR Pfam: PF00089; trypsin; 1.
DR Pfam: PF00092; vwa; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00453; VWFADOMAIN.
DR SMART: SM00332; CCP; 3.
DR SMART: SM00620; Tryp_SPC; 1.
DR SMART: SM00327; VWA; 1.
DR PROSITE: PS50240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
Query Match 5.8%; Score 83; DB 1; Length 764;
Best Local Similarity 24.1%; Pred. No. 12;
Matches 49; Conservative 21; Mismatches 71; Indels 62; Gaps 12;
QY 24 GTGLYPMRCPPKNLALLPFLPLGGG-----GSGSEKSV----- 60
DE 2 GSNLSP-----QLCLMPFILLGSGGVTTTPWSLARFQSGCSLEGVEIKGSRLLQEG 55
QY 61 SKMAAAPSG--PSAPEAVTARLVGLVFWVSVTGFWGAVATS-----AGGEESLK-- 110
DB 56 QALEVVCSGGYPPVQVTRICR-----STGSWILKTDQOKTVRKACRAHCPR 105
QY 111 -EDLVGVYICKPKINDATQEPVNC-TNYTAHVSCTFPAPN-TCKDSS--QNFHTFTGNE 166
DB 106 PHDFENGVEWPRSPYVNVSDSIFHCYDGYTLRGS-----NRTQVNGRWSGQTACDNG 161
QY 167 VGFFK----PISCRNVNGYSYKV 185
DB 162 AGYCSNPGIPGTRKV-GSQYRL 183
RESULT 10
LAMP_HUMAN
ID LAMP_HUMAN STANDARD; PRT; 338 AA.
AC Q13449;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Limbic system-associated membrane protein precursor (LSAMP).
GN LSAMP OR LAMP.
OS Homo sapiens (Human).
OC Fukuyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-96235133; PubMed-8666243;
RA Pimente A.F., Fischer I., Levitt P.;
RT "cDNA cloning and structural analysis of the human limbic-system-
RT associated membrane protein (LAMP).";
RL Gene 170:188-195(1996).
CC -!- FUNCTION: MEDIATES SELECTIVE NEURONAL GROWTH AND AXON TARGETING.
CC CONTRIBUTES TO THE GUIDANCE OF DEVELOPING AXONS AND REMODELING OF
CC MATURE CIRCUITS IN THE LIMBIC SYSTEM. ESSENTIAL FOR NORMAL GROWTH
CC OF THE HYPOCAMPAL MOSSY FIBER PROJECTION (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- TISSUE SPECIFICITY: EXPRESSED ON LIMBIC NEURONS AND FIBER TRACTS
CC AS WELL AS IN SINGLE LAYERS OF THE SUPERIOR COLLICULUS, SPINAL
CC CHORD AND CEREBELLUM.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. IGLON
CC SUBFAMILY.
CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U41901; AAC50569.1; .
DR PIR: J04776; J04776.
DR Genbank: HGRC:6705; LSAMP.
DR MIM: 603241; .
DR GO: GO:007399; P:neurogenesis; TAS.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003598; Ig-C2.
DR Pfam: PF00047; Ig_3.
DR SMART: SM00408; IGC2; 2.
DR PROSITE: PS00835; IG-LIKE; 3.
KW Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
KW Repeat; Signal.
FT SIGNAL 1 28
FT CHAIN 29 315
FT PROPEP 316 338
FT DOMAIN 29 122
FT DOMAIN 132 214
FT DOMAIN 219 304
FT DISULFID 53 111
FT DISULFID 153 197
FT DISULFID 239 290
FT CARBOHYD 40 40
FT CARBOHYD 66 66
FT CARBOHYD 136 136
FT CARBOHYD 148 148
FT CARBOHYD 279 279
FT CARBOHYD 287 287
FT CARBOHYD 300 300
FT CARBOHYD 315 315
FT LIPID 315 315
SQ SEQUENCE 338 AA: 37308 MW: 034552860DFD092F CRC64;
Query Match 5.7%; Score 82.5; DB 1; Length 338;
Best Local Similarity 29.6%; Pred. No. 5.2;
Matches 37; Conservative 14; Mismatches 47; Indels 27; Gaps 7;
QY 101 SAGGESLKCEDLVG-----QYCKDPKINDATQPVNCTNTHAVSCFPAPN:TCQDSS 156
DB 230 TTGROASLKCEASAVPAPDFEYRDDTINSANGLEIKSTE---GQSSLTYNVT-EEHY 235
QY 157 GNETHFTGNEVG-----FFKPIPCRNWNGSYKVAVALSLFLGLGADRFY-GYPALG 209
DB 286 GNYTCVAANKLGVTNASLVIFRPGSVRGING-SISLAVPL-----WL-----LAASLAC 333
QY 210 LKFC 214
DB 334 LLSKC 338
RESULT 11
VG32-BPM02
ID VG32-BPM02 STANDARD: PRT: 223 AA.
AC 064226;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gene 32 protein (GP32).
GN 32.
OS Mycobacteriophage D29.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=28369;
RN [1]
RP SEQUENCE FROM N.A.
RX FORD M.E., Sarkis G.J., Belanger A.E., Hendrix R.W., Hatfull G.F.;
RA "Genome structure of mycobacteriophage D29: implications for phage
RT

RT evolution.";
RE J. Mol. Biol. 279:143-164(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF022214; AAC18473.1; .
DR PIR: F72803; F72803.
SQ SEQUENCE 223 AA: 21822 MW: 33CD0DC310038AD4 CRC64;
Query Match 5.6%; Score 80.5; DB 1; Length 223;
Best Local Similarity 30.7%; Pred. No. 4.8;
Matches 27; Conservative 8; Mismatches 36; Indels 17; Gaps 1;
QY 29 PMRGPFKNLALLPFLP-----LLGGSGSGSEKVSVMKMAAWPSGSPAPAV-A 79
DB 37 PVLTPVTAVGAYTYNIPAOAEFIDVLLGAGGGGQ-----MGSAWAGGGGFGGSWV-A 91
QY 80 RL---VGVLFVSVYTGPCWCAVATSAGG 104
DB 92 TLRRGVDIPMAVTQITGVIGAGGTAGPG 119
RESULT 12
ENT1_HUMAN
ID ENT1_HUMAN STANDARD: PRT: 455 AA.
AC Q99808;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Equilibrative nucleoside transporter 1 (Equilibrative
DE nitrobenzylmercaptopurine riboside-sensitive nucleoside transporter)
DE (Equilibrative NBMPR-sensitive nucleoside transporter) (Nucleoside
DE transporter, es-type).
DE SLC29A1 OR ENT1.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-21.
RX TISSUP-Placenta;
MEDLINE-97140266; PubMed-8986748;
RA Griffiths M., Beaumont N., Yao S.Y.M., Sundaram M., Bouman C.E.,
RA Davies A., Kwong F.Y.P., Coe I., Cass C.E., Young J.D., Baldwin S.A.;
RA "Cloning of a human nucleoside transporter implicated in the cellular
RA uptake of adenosine and chemotherapeutic drugs.";
RA Nat. Med. 3:89-93(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUP-Jejunum, and Small intestine;
RA Lum P.Y., Ngo L.Y., Bakken A.H., Unadkat J.D.;
RA "Critical structural determinants for high affinity binding of
RA nucleosides to the equilibrative NBMPR-sensitive nucleoside
RA transporter (es) cloned from the human jejunum.";
RA Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MEDIATES BOTH INFUX AND EFFLUX OF NUCLEOSIDES ACROSS
CC THE MEMBRANE (EQUILIBRATIVE TRANSPORTER). IT IS SENSITIVE (ES) TO
CC LOW CONCENTRATIONS OF THE INHIBITOR NITROBENZYLMECAPTOPURINE
CC RIBOSIDE (NBMPR) AND IS SODIUM- INDEPENDENT. IT HAS A HIGHER
CC AFFINITY FOR ADENOSINE. INHIBITED BY DIPYRIDAMOLE AND DILAZEP
CC (ANTICANCER CHEMOTHERAPEUTICS DRUGS).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART, BRAIN, MAMMARY GLAND,
CC ERYTHROCYTES AND PLACENTA, AND ALSO IN FETAL LIVER AND SPLEEN.
CC -1- PTM: GLYCOSYLATED.
CC -1- SIMILARITY: BELONGS TO THE SLC29A FAMILY OF TRANSPORTERS.
CC -----


```

CC -!- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a
CC fatty acid anion.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE ARHYDROLASE SUPERFAMILY. LIPASE FAMILY.
CC -!- SIMILARITY: Contains 1 PLAT domain.
CC -!- DATABASE: NAME-Worthington enzyme manual;
CC WWW="http://www.worthington-biochem.com/manual/L/PL.html".
CC PDB: 1FTH: 07-DEC-96.
DR InterPro: IPRO00734; Lipase.
DR InterPro: IPRO01024; Lipoxigenase_LH2.
DR InterPro: IPRO00379; Ser_cstrs_site.
DR Pfam: PF00151; Lipase; 1.
DR Pfam: PF01477; PLAT; 1.
DR PRINTS: PR00821; TAGLIPASE.
DR SMART: SM00308; LH2; 1.
DR PROSITE: PS00120; LIPASE_SER; 1.
DR PROSITE: PS00095; PLAT; 1.
KW Hydrolase; Lipid degradation; Pancreas; Glycoprotein; 3D-structure.
FT DOMAIN 339 450
FT ACT_SITE 153 153 CHARGE RELAY SYSTEM.
FT ACT_SITE 177 177 CHARGE RELAY SYSTEM.
FT ACT_SITE 264 264 CHARGE RELAY SYSTEM.
FT DISULFID 4 10
FT DISULFID 91 102 IN ISOMER 1.
FT DISULFID 91 104 IN ISOMER 2.
FT DISULFID 238 262
FT DISULFID 286 297
FT DISULFID 300 305
FT DISULFID 434 450
FT CARBOHYD 167 167 N-LINKED (GLCNAC...)
SQ SEQUENCE 450 AA; 50084 MW; 76E13BB3B451PCE CRC64;

Query Match 5.6%; Score 80; DB 1; Length 450;
Best Local Similarity 20.2%; Pred. No. 12;
Matches 47; Conservative 30; Mismatches 96; Indels 60; Gaps 10;

QY 1 MHILKSPNVIP-----RAHGQKNTRRDQ-----TCLYPMRGPFKNLALL---40
DB 133 VEVLKSSLGSPSNVHVIGHSLGSHAAGEARTNGTIERITCLDPAEPCFGGTPE:VRL 192
QY 41 -----PFSILPLGGSGGSEKVSXKMAAWPSG-----PSAPEAVTARLVGV 34
DB 193 DPSDAKFDVVIHTDAAPIINLGFMSQTVGHLDP---PPNGGKOMPCCOKNLSOIVEI 244
QY 85 LWFVSVTTPGWAVATSAGGESLCEDLKVGQYICKDKPKINDATQEPVNCINYARHS- 143
DB 250 -----DGIW-----EGTRQFVACNHLSRYKYA-CSILAPDGGAGPGCDSYNVFTAN 295
QY 144 -GFPAPNITCKDSSGNETHFTG-----NEVGFFKP-SCHNVNGYSYKVAVALS 190
DB 296 KCFPCSECGPQMGHVADRFPGKTNVGSQVYINTGFSNFAWRWYKVSVTLS 348

RESULT 15
OARI_LYMNST
ID OARI_LYMNST STANDARD; PRT: 638 AA.
AC 077408;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Octopamine receptor 1 (OAR).
OS Lymnaea stagnalis (Great pond snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC Lymnaeidae; Lymnaeinae; Lymnaea.
OX NCBI_Taxid=6523;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC TISSUE=ONS;
RX MEDLINE=97347296; PubMed=9203635;
RA Gerhardt C.C., Baker R.A., Plek R.J., Planta R.J., Vreugdenhil E.,
RA Leysen J.E., van Heerikhuizen H.;
RT *Molecular cloning and pharmacological characterization of a molluscan
RT octopamine receptor*;
```

```

RL Mol. Pharmacol. 51:293-300(1997).
CC -!- FUNCTION: G-PROTEIN COUPLED RECEPTOR FOR OCTOPAMINE (OA), WHICH IS
CC A NEUROTRANSMITTER, NEUROHORMONE, AND NEUROMODULATOR IN
CC INVERTEBRATES. ACTIVATION OF THIS RECEPTOR BY OCTOPAMINE INDUCES
CC AN INCREASE IN BOTH INOSITOL PHOSPHATES AND CYCLIC AMP. THE
CC COUPLING TO ADENYL CYCLASE SEEMS TO BE LESS EFFICIENT THAN THE
CC COUPLING TO PHOSPHOLIPASE C. THE RANK ORDER OF POTENCY FOR
CC AGONISTS IS P-SYNEPHRINE >= CLONIDINE > P-OCTOPAMINE =
CC XYLOMETHAZOLINE = PHENYLEPHRINE = OXYMETAZOLINE > B-HT920 >
CC SERTONIN = P-TYRAMINE > EPINEPHRINE > NOREPINEPHRINE >
CC METHOXAMINE = DOPAMINE = HISTAMINE. FOR ANTAGONISTS, THE RANK
CC ORDER IS YOHIMBINE > CHLOPRAMAZINE / SPIPERONE > PHENTOLAMINE >
CC MANSERINE > RAUWOOLSCINE > PRAZOSIN > ALPRENOLOL / PROPRANOLOL >
CC PINDOLOL.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE CENTRAL NERVOUS SYSTEM.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U62771; AAC61296.1; -.
CC HSSP: 229274; 1MM5;
CC InterPro: IPR000276; GPCR_Rhodpsn.
CC Pfam: PF00001; 7tm1.1;
CC PRINTS: PR00237; GPCR_RHODPSN.
CC PROSITE: PS00237; G-PROTEIN RECF_FL_1; 1.
CC PROSITE: PS0262; G-PROTEIN RECF_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 28
FT TRANSMEM 29 53 POTENTIAL.
FT DOMAIN 54 64 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 65 87 POTENTIAL.
FT DOMAIN 88 102 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 103 124 POTENTIAL.
FT DOMAIN 125 147 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 148 167 POTENTIAL.
FT DOMAIN 168 239 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 240 259 POTENTIAL.
FT DOMAIN 260 520 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 521 545 POTENTIAL.
FT DOMAIN 546 551 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 552 575 POTENTIAL.
FT DOMAIN 576 638 CYTOPLASMIC (POTENTIAL).
FT DISULFID 101 230 BY SIMILARITY.
FT CARBOHYD 178 178 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 207 207 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 215 215 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 638 AA; 70000 MW; 65FA928B5C01D34F CRC64;

Query Match 5.6%; Score 80; DB 1; Length 638;
Best Local Similarity 24.8%; Pred. No. 18;
Matches 31; Conservative 18; Mismatches 56; Indels 20; Gaps 4;

QY 149 NITCKD-----SSGNETHFTGNEVGFFKPISCHNVNGYSYKVAVALSFLGLWGA 198
DB 69 NLACADLLGLVLVLPFSVAVNEI---KDWIFGHVWCQVWLAVDVMVCTASILNCCISL 124
QY 199 DRFY-----LGPALGLLKFCVTGFCGIGCSLIDFLISQVIVGSPSGSYIDYGTGLT 253
DB 125 DYLAITPIRYPGLMSAKRAKTLVAGVM-LFSFVICPPLIGMNDGGDGYNGTTAT 183
QY 254 RLSIT 258
DB 184 PIPVT 188

Search completed: September 26, 2003, 17:25:45
```


XX Bard JA, Jacobsen JS, Kajkowski EM, Ozenberger BA;
PI Walker SG;

XX WPI: 1999-080736/27.
 DR N-PSDB; AAX05735.
 XX
 PT Polynucleotide encoding beta-amyloid peptide binding protein - used
 PT to identify inhibitors of beta-amyloid peptide for treating
 PT Alzheimer's disease
 XX
 PS Claim 7; Pages 43-44; 59pp; English.
 CC
 CC The present sequence represents a beta-amyloid peptide binding protein
 CC (BBP). The polynucleotide comprising the entire BBP nucleotide sequence
 CC of clone BBP1-f1 is deposited under the accession number: AACC 94517. The
 CC polynucleotide comprising a fragment of BBP (nucleotides 202-807 of the
 CC full length BBP) of clone PK196 is deposited as ATCC 98399. Host cells
 CC transformed with a vector comprising the BBP nucleic acid are used for
 CC the recombinant production of the protein. The protein can be used in a
 CC method for diagnosing a disease characterised by aberrant expression of
 CC human beta-amyloid protein (BAP). The protein can also be used in a
 CC method for screening for compounds which regulate expression of a BAP
 CC binding protein. The proteins, antibodies and identified compounds can be
 CC used in the treatment or prevention of Alzheimer's disease.
 XX
 SQ Sequence 269 AA:
 Query Match 100.0%; Score 33; DB 20; Length 269;
 Best Local Similarity 100.0%; Pred. No. 3.9e-24;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VAVALSFLGWLGDRLFYLGYPALGLLKFTCTVG 33
 Db 185 VAVALSFLGWLGDRLFYLGYPALGLLKFTCTVG 217
 RESULT 2
 AAY70759
 ID AAY70759 standard; Protein: 269 AA.
 AC AAY70759;
 XX
 DT 24-JUL-2000 (first entry)
 DE Human beta-amyloid peptide (BAP) binding protein, BBP1.
 XX
 KW Beta-amyloid peptide binding protein; BAP; BAP; tumour; suppressor;
 KW G-protein coupled receptor; GPCR; integral membrane protein; antigen;
 KW neuronal cell; nonhuman primate; NHP; G-protein signalling pathway;
 KW apoptosis; immunogen; therapeutic; treatment; prevention; diagnostic.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 177..198
 FT /label= Transmembrane_domain_1
 FT 199..201
 FT /label= DRF_motif
 FT /note= "Substitution of the Arg abrogates protection"
 FT 213..238
 FT /label= Transmembrane_domain_2
 XX
 PN WO2000022125-A2.
 XX
 PD 20-APR-2000.
 XX
 PF 13-OCT-1999; 99WO-US21621.
 XX
 PR 13-OCT-1998; 98US-0104104.
 XX
 PA (AMHP) AMERICAN HOME PROD CORP.
 XX
 XX Ozenberger BA, Kajkowski EM, Lo CF;
 PI
 DR WPI: 2000-317982/27.

DR N-PSDB; AAZ52369.
 XX
 PT Novel G-protein-coupled receptor-like proteins and polynucleotides
 PT useful for regulating apoptosis, comprises integral membrane protein
 PT traversing the membrane twice.
 XX
 PS Example 1; Page 62-63; 68pp; English.
 XX
 CC The present sequence is the beta-amyloid peptide (BAP) binding protein-1
 CC (BBP1). It is an integral membrane protein, that traverse the membrane
 CC twice. It is related to G protein-coupled receptor (GPCR) protein
 CC superfamily. It interacts with G-alpha proteins and regulates the
 CC activity of G-protein signalling pathways. BBP genes are widely expressed
 CC in neuronal cells of nonhuman primate (NHP) brain and overexpressed in
 CC some tumours. It functions as a suppressor of apoptosis induction. BBP
 CC proteins are used as immunogens to raise antibodies, useful as
 CC therapeutics and as antigens in solid phase assays. They are also useful
 CC as reagents to identify molecules which effect the interaction of BBP and
 CC a cloned protein, that are useful in the treatment or prevention of
 CC diseases associated with apoptosis. The polynucleotides are useful for
 CC diagnostics.
 CC Note: In claim 5, the patent claims an amino acid sequence from figure 2.
 CC However, figure 2 does not contain any sequence. It is inferred from the
 CC disclosure that the figure 2 sequence refers to BBP1 protein, shown in
 CC this sequence.
 XX
 SQ Sequence 269 AA:
 Query Match 100.0%; Score 33; DB 21; Length 269;
 Best Local Similarity 100.0%; Pred. No. 3.9e-24;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VAVALSFLGWLGDRLFYLGYPALGLLKFTCTVG 33
 Db 185 VAVALSFLGWLGDRLFYLGYPALGLLKFTCTVG 217
 RESULT 3
 AAE33877
 ID AAE33877 standard; Protein: 269 AA.
 XX
 AC AAE33877;
 XX
 DT 02-MAY-2003 (first entry)
 DE Human: BBP-1 protein.
 XX
 KW Human: beta-amyloid peptide-binding protein; BAP; Abeta; betaAP; BBP;
 KW Alzheimer's disease; AD; transgenic; transgenic animal; gene therapy;
 KW neuroprotective; neurotropic.
 XX
 OS Homo sapiens.
 XX
 FN WO200290499-A2.
 XX
 PD 14-NOV-2002.
 XX
 PF 06-MAY-2002; 2002WO-US14223.
 XX
 PR 09-MAY-2001; 2001US-0852100.
 XX
 PA (AMHP) WYETH.
 XX
 PI Ozenberger BA, Bard JA, Kajkowski EM, Jacobsen JS, Walker SG;
 PT Sofia HJ, Howland DS;
 XX
 DR WPI: 2003-120537/11.
 DR N-PSDB; AAD51940.
 XX
 XX New human beta-amyloid peptide-binding protein, useful for diagnosing
 PT and/or treating diseases associated with aberrant expression of
 PT beta-amyloid peptide, e.g. Alzheimer's disease.
 XX

PS Claim 4; Page 84-85; 85pp; English.

XX The present invention relates to novel human beta-amyloid peptide (BAP;
CC Abeta, betaAPP)-binding (BBP) proteins and polynucleotides encoding such
CC proteins. BBP sequences are useful to diagnose and/or treat diseases
CC associated with aberrant expression of human BAP such as Alzheimer's
CC disease (AD). They are used to generate transgenic animals. Sequences
CC of the invention are also used in gene therapy. The present sequence
CC is human BBP-1 protein.

XX Sequence 269 AA;

Query Match 100.0%; Score 33; DB 24; Length 269;

Best Local Similarity 100.0%; Pred. No. 3; 9e-24;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAVALSLFLGWLGLADREYLGYPALGLLKFCTVG 33

DB 185 VAVALSLFLGWLGLADREYLGYPALGLLKFCTVG 217

RESULT 4

AAV12358

ID AAV12358 standard; Protein; 139 AA.

XX AC AAV12358;

XX 17-JUN-1999 (first entry)

DI Human 5' EST secreted protein SEQ ID NO:389.

DE Human; secreted protein; EST; expressed sequence tag; diagnosis;
XX forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.

OS Homo sapiens.

XX W09906548-A2.

XX 11-FEB-1999.

XX 31-JUL-1998; 98WO-IB01222.

XX 01-AUG-1997; 97US-0905135.

XX (GEST) GENSET.

XX Duclert A, Dumas Milne Edwards J, Lacroix B;

XX WPI: 1999-153778/13.

XX N-PSDB: AAX41191.

XX New nucleic acids encoding human secreted proteins - obtained from
XX cDNA libraries prepared from e.g. liver, ovary, brain, prostate,
XX kidney, lung, umbilical cord, placenta and colon tissue

XX Claim 27; Page 714-715; 824pp; English.

XX AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins, and encode the proteins given in AAX12261 to
CC AAX12514, respectively. The proteins given represent the signal peptide
CC and an N-terminal fragment of a secreted protein. The nucleic acid
CC sequences can be used for producing secreted human gene products. They
CC can also be used to develop products for diagnosis and therapy. The
CC proteins obtained may have cytokine activity, cell
CC proliferation/differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products

CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptide can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell.

XX Sequence 139 AA;

Query Match 51.5%; Score 17; DB 20; Length 139;

Best Local Similarity 100.0%; Pred. No. 6; 9e-09;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAVALSLFLGWLGLADRF 17

DB 123 VAVALSLFLGWLGLADRF 139

RESULT 5

ABR11574

ID ABR11574 standard; peptide; 225 AA.

XX AC ABR11574;

XX 11-JAN-2002 (first entry)

XX Human secreted protein homologue, SEQ ID NO:1944.

XX Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; anti-inflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnerary; antiulcer.

XX Homo sapiens.

XX W0200157188-A2.

XX 09-AUG-2001.

XX 03-FEB-2001; 2001WO-US03800.

XX 03-FEB-2000; 2000US-C496914.

XX 27-APR-2000; 2000US-C0560875.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI: 2001-457740/49.

XX N-PSDB: ABA06818.

XX Human proteins and DNA encoding sequences useful for preventing,
XX treating or ameliorating a medical condition in a mammalian subject
XX e.g. arthritis and cancer -

XX Claim 20; Page 214; 1963pp; English.

XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC invention also relates to vectors and recombinant host cells comprising
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotide
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby

CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention.

XX Sequence 225 AA;

Query Match 27.3%; Score 9; DB 22; Length 225;
 Best Local Similarity 100.0%; Pred. No. 0.62;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 12 LGADRFYLG 20
 DB 176 LGADRFYLG 184
 |||||

RESULT 6
 ABB77688
 ID ABB77688 standard; peptide: 24 AA.

XX AC ABB77688;

XX DT 01-JUL-2002 (first entry)

XX DE New peptide vector#4.

XX Intracellular delivery; transfection agent; cancer; infectious disease;
 KW peptide vector.

XX OS Synthetic.

XX PN WO200210201-A2.

XX PD 07-FEB-2002.

XX PF 26-JUL-2001; 2001WO-US234C6.

XX PR 31-JUL-2000; 2000US-221932P.

XX PA (ACTI-) ACTIVE MOTIF.

XX PA (CNRS) CENT NAT RECH SCI.

PI Divida G, Morris M, Mery J, Heitz F, Fernandez J, Archdeacon J;
 PI Horndorp K;

XX WPI; 2002-329441/36.

XX Transfection agent that comprises a peptide comprising hydrophobic and

PT hydrophilic domain and having amino acid residues of specified length
 PT is useful for a non-covalent association with and transport of a
 PT heterologous compound into a cell

XX Example 2; Page 61; 156pp; English.

XX The invention relates to a transfection agent comprises a peptide of
 XX about 16 - 30 amino acids in length. Peptides of the invention comprise
 XX a hydrophobic domain, a hydrophilic domain, optionally a spacer sequence
 XX between the domains and a functional group conjugated to at least one
 XX terminal of the peptide. Peptides of the invention are useful for a
 XX non-covalent association with and transport of a heterologous compound
 XX into a cell. They are also useful for promoting the cellular
 XX internalisation of at least one member e.g. peptide, proteins,
 XX antibodies, their derivatives and/or conjugates. They may form part of a
 XX pharmaceutical composition to deliver the compound selected from a
 XX diagnostic or therapeutic compound, to treat at least one condition such
 XX as cancer or an infectious disease, or which targets a cancerous cell or
 XX pathogen-infected cell and to deliver a peptide or inhibitor that
 XX disrupts the activity of the enzyme. The agent of the invention has a
 XX transfection efficiency of at least 5% for at least two of the members of
 XX the group of the compounds. The agent has a good delivery efficiency for
 XX a broad spectrum of compounds and cell types, has a low toxicity, are
 XX easy to handle and easy to formulate in conjunction with the many
 XX different compound types that it can deliver. The peptides are serum
 XX sensitive, thus they bind particularly well for systemic and/or localised
 XX in patients. The current sequence represents a new amphipathic peptide
 XX vector of the invention that contains a cationic nuclear localisation
 XX sequence.

XX Sequence 24 AA;

Query Match 24.2%; Score 8; DB 23; Length 24;
 Best Local Similarity 100.0%; Pred. No. 0.71;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LFLGLWGA 14
 DB 3 LFLGLWGA 10
 |||||

RESULT 7
 ABB81176

ID ABB81176 standard; peptide: 27 AA.

XX AC ABB81176;

XX DT 25-NOV-2002 (first entry)

XX DE Signal sequence based peptide I MPS peptide.

XX Red blood cell; RBC; virus; vertebrate; therapeutic; tissue imaging;
 KW gp4.; membrane translocation.

XX Human immunodeficiency virus type 1.

XX Kitesus macaque polyoma virus.

XX PN WO200260416-A1.

XX PD 08-AUG-2002.

XX PF 01-FEB-2002; 2002WO-GB00437.

XX PR 01-FEB-2001; 2001GB-0G02561.

XX PR 16-FEB-2001; 2001US-269528P.

XX PA (GEND-) GENDEL LTD.

XX PI Mchale AP, Craig R;

XX WPI; 2002-643355/69.

XX Delivering agent to target site in vertebrate comprises loading red

PT blood cell with virus or virus-like particle comprising agent,
 PI sensitizing cell, introducing cell into vertebrate and applying energy
 XX to release virus particle from cell -
 XX
 PS Disclosure: Page 55; 87pp; English.
 XX
 CC The invention relates to delivering an agent to a target site in a
 CC vertebrate. The method involves (a) loading a red blood cell (RBC) with a
 CC virus or a virus-like particle (I) comprising an agent; (b) sensitizing
 CC RBC to render it more susceptible to disruption than unsensitized RBC;
 CC (c) introducing RBC into a vertebrate, and (d) applying energy to release
 CC (I). Steps (a) and (b) may be performed in any order. (II) (RBC loaded
 CC with a virus or a virus-like particle comprising a therapeutic agent) is
 CC useful for the delivery of a therapeutic agent to a target site in a
 CC vertebrate, or in the preparation of a medicament for delivery of a
 CC therapeutic agent to a target site in a vertebrate. (II) is also useful
 CC for delivering one or more agents to a vertebrate and for treating or
 CC preventing a disease. The method is useful for delivering agents such as
 CC those useful for imaging of tissues in vivo or ex vivo, preferably for
 CC delivering an agent to a subcellular organelle such as nucleus,
 CC mitochondria, Golgi or endoplasmic reticulum. The present sequence
 CC represents a signal sequence based peptide: WPS peptide, a chimera of
 CC the hydrophobic terminal domain of viral gp41 protein and the nuclear
 CC localisation signal (NLS) from the SV40 large antigen. This fragment has
 CC been found to be active in membrane translocation.
 XX
 SQ Sequence 27 AA:

Query Match 24.2%; Score 8; DB 23; Length 27;
 Best Local Similarity 100.0%; Pred. No. 0.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LFLGWLGA 14
 DQ 3 LFLGWLGA 1C
 |||||

RESULT 8
 ABG78990
 ID ABG78990 standard; Peptide: 27 AA.

XX
 AC ABG78990;

XX 15-NOV-2002 (first entry)

XX Cell penetrating peptide Signal-peptide II.

XX Cell penetrating peptide; cancer; tumour; melanoma; thymoma;
 KW lymphoma; sarcoma; lung cancer; non-Hodgkin's lymphoma; leukemia;
 KW Hodgkin's lymphoma; uterine cancer; cervical cancer; bladder cancer;
 KW kidney cancer; adenocarcinoma; breast cancer; prostate cancer;
 KW ovarian cancer; pancreatic cancer; vaccine; dendritic cell;
 KW tumour infiltrating lymphocyte; TIL; human leukocyte antigen: HLA;
 KW cytostatic.

XX Unidentified.

XX WO200264057-A2.

XX 22-AUG-2002.

XX 15-FEB-2002; 2002WO-US05212.

XX 15-FEB-2001; 2001US-268687P.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Wang R;

XX WPI: 2002-627577/67.

XX Novel composition for treating a disease in an animal, comprises an
 PT immune effector cell and cell penetrating peptide associated with an

PT antigen or antibody -
 XX
 PS Disclosure: Page 11; 61pp; English.
 XX
 CC The invention relates to a composition (I) comprising an immune effector
 CC cell and a cell penetrating peptide (CPP) associated with an antigen or
 CC antibody. Also included are (1) a vaccine comprising (I), CPP
 CC associated with an antigen, and a pharmaceutically acceptable carrier
 CC and (2) preparing a composition for a disease, by providing (I)
 CC and CPP associated with an antigen for disease, and introducing the
 CC antigen-associated CPP to (I), where antigen enters into the cell.
 CC The antigens are, for example, tumour antigen derived epitopes
 CC recognised by tumour infiltrating lymphocytes (TIL) of HLA (human
 CC leukocyte antigen) class I or II. The composition is useful for enhancing
 CC immunity in an animal to a disease, by administering a mature dendritic
 CC cell comprising CPP associated with an antigen to disease, to the animal,
 CC such that following the administration, animal is protected from disease,
 CC where the animal comprises both CD4+ and CD8+ T cells. It is also useful
 CC for treating a disease (e.g. cancer, tumour, melanoma, thymoma,
 CC lymphoma, sarcoma, lung cancer, non-Hodgkin's lymphoma, leukemia,
 CC Hodgkin's lymphoma, uterine cancer, cervical cancer, bladder cancer,
 CC kidney cancer, adenocarcinoma, breast cancer, prostate cancer,
 CC ovarian cancer and pancreatic cancer). The animal is further subjected to
 CC a cancer treatment including surgery, radiation, chemotherapy or gene
 CC therapy. The administration of (I), preferably dendritic cell is prior
 CC to, subsequent to or concurrent with, the cancer treatment. The present
 CC sequence is cell penetrating peptide of the invention.

XX Sequence 27 AA:

Query Match 24.2%; Score 8; DB 23; Length 27;
 Best Local Similarity 100.0%; Pred. No. 0.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LFLGWLGA 14
 DQ 3 LFLGWLGA 10
 |||||

RESULT 9
 AAE23685

XX AAE23685 standard; peptide: 27 AA.

XX AAE23685;

XX 10-SEP-2002 (first entry)

XX Fluorescently labelled RNA binding peptide #2.

XX RNA binding protein; mRNA quantification; gene expression.

XX Unidentified.

XX WO200227031-A2.

XX 04-APR-2002.

XX 28-SEP-2001; 2001WO-US0438.

XX 28-SEP-2000; 2000US-236407P.

XX (CELL-) CELLWIGS INC.

XX Busa WB;

XX WPI: 2002-452280/48.

XX Quantifying target gene expression in living cells that possess a
 PT target gene of interest tagged with the binding site for an RNA binding
 PT protein and fluorescently labeled RNA binding polypeptide including an
 PT RNA binding domain -
 XX
 PS Claim 45; Page 44; 51pp; English.

XX The present invention relates to a method of quantifying the expression
 CC of target genes in living cells. The method involves providing cells that
 CC possess a target gene of interest which has been tagged with the binding
 CC site for an RNA binding protein and a fluorescently labelled RNA binding
 CC polypeptide that includes an RNA binding domain and calculating the
 CC quantity of target gene expression in the cells using fluorescence
 CC signalling techniques. The method is useful for quantifying expression
 CC of one or more target genes in living cells which comprise two or more
 CC distinct populations of cells. It is used to quantitate the expression
 CC of any target gene, including expression of protein-encoding messenger
 CC RNA genes, ribosomal RNA encoding genes and transfer RNA encoding genes
 CC so long as the RNA expression product from the target gene possesses a
 CC sequence or structure (the RNA tag) that is bound specifically by the
 CC RNA binding polypeptide being used. The present sequence is a
 CC fluorescently labelled RNA binding peptide.

XX Sequence 27 AA;
 SO

Query Match 24.2%; Score 8; DB 23; Length 27;
 Best Local Similarity 100.0%; Pred. No. 0.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LFLGWLGA 14
 Db 3 LFLGWLGA 10
 |||||
 11111111

RESULT 10
 ABB77687
 ID ABB77687 standard; peptide: 27 AA.
 XX AC ABB77687;
 XX DT 01-JUL-2002 (first entry)
 XX DE New peptide vector#3.
 XX KW Intracellular delivery; transfection agent; cancer; infectious disease;
 XX peptide vector.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Misc-difference 7
 FT /note- "residue may be substituted with Phe"
 FT Misc-difference 23
 FT /note- "residue may be substituted with Ser"
 XX WO200210201-A2.
 XX PN 07-FEB-2002.
 XX PD 26-JUL-2001; 2001WO-US23406.
 XX PF 31-JUL-2000; 2000US-221932P.
 XX PR (ACT1-) ACTIVE MOTIF.
 XX (CNRS) CENT NAT RECH SCI.
 XX PA Divida G, Morris M, Mery J, Heitz F, Fernandez J, Archdeacon J;
 XX Horndorp K;
 XX WIPI; 2002-329441/36.
 XX DR Transfection agent that comprises a peptide comprising hydrophobic and
 PT hydrophilic domain and having amino acid residues of specified length.
 PT is useful for a non-covalent association with and transport of a
 PT heterologous compound into a cell -
 XX Example 2; Page 61; 156pp; English.
 XX PS The invention relates to a transfection agent comprises a peptide of

CC about 16 - 30 amino acids in length. Peptides of the invention comprise
 CC a hydrophobic domain, a hydrophilic domain, optionally a spacer sequence
 CC between the domains and a functional group conjugated to at least one
 CC terminal of the peptide. Peptides of the invention are useful for a
 CC non-covalent association with and transport of a heterologous compound
 CC into a cell. They are also useful for promoting the cellular
 CC internalisation of at least one member e.g. peptide, proteins,
 CC antibodies, their derivatives and/or conjugates. They may form part of a
 CC pharmaceutical composition to deliver the compound selected from a
 CC diagnostic or therapeutic compound, to treat at least one condition such
 CC as cancer or an infectious disease, or which targets a cancerous cell or
 CC pathogen-infected cell and to deliver a peptide or inhibitor that
 CC disrupts the activity of the enzyme. The agent of the invention has a
 CC transfection efficiency of at least 5% for at least two of the members of
 CC the group of the compounds. The agent has a good delivery efficiency for
 CC a broad spectrum of compounds and cell types, has a low toxicity, are
 CC easy to handle and easy to formulate in conjunction with the many
 CC different compound types that it can deliver. The peptides are serum
 CC sensitive, thus they bode particularly well for systemic and/or localised
 CC in patients. The current sequence represents a new amphipathic peptide
 CC vector of the invention that contains a cationic nuclear localisation
 CC sequence.

XX Sequence 27 AA;
 SO

Query Match 24.2%; Score 8; DB 23; Length 27;
 Best Local Similarity 100.0%; Pred. No. 0.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LFLGWLGA 14
 Db 3 LFLGWLGA 10
 |||||
 11111111

RESULT 11
 AA076348
 ID AA076348 standard; Peptide: 27 AA.
 XX AC AA076348;
 XX DT 18-JUN-2002 (first entry)
 XX DE Signal sequence based peptide I.
 XX KW Membrane translocation signal; signal sequence based peptide 1;
 KW red blood cell vehicle; polypeptide delivery; viral gp41 protein;
 KW simian virus 40; SV40; large antigen.
 XX OS Rhesus macaque polyoma virus.
 XX Synthetic.
 XX PN WO200207752-A2.
 XX PD 31-JAN-2002.
 XX PF 24-JUL-2001; 2001WO-GB03327.
 XX PR 24-JUL-2000; 2000WO-GB02848.
 XX PR 09-AUG-2000; 2000WO-GB03056.
 XX PR 01-FEB-2001; 2001WO-GB00417.
 XX PR 16-FEB-2001; 2001US-0785802.
 XX FA (GEND-) GENDEL LTD.
 XX PI Craig R;
 XX WIPI; 2002-280593/32.
 XX DR Preparing a red blood cell vehicle suitable for delivering an agent to
 PT a target site in a vertebrate due to loading the red blood cell with an
 PT agent-membrane translocation sequence -
 XX Example 2; Page 44; 135pp; English.
 XX PS Disclosure; Page 44; 135pp; English.

XX The invention describes a method of preparing a red blood cell vehicle
 CC suitable for delivering an agent to a target site in a vertebrate
 CC comprising providing a red blood cell and loading the red blood cell with
 CC an agent-MTS (membrane translocation sequence) conjugate. The red blood
 CC cells produced may be used in the preparation of a medicament for
 CC delivery of an agent to or at a target site and of one or more agents to
 CC a vertebrate. The agent is actively released from the red blood cell
 CC vehicle by application of a stimulus to disrupt the red blood cell
 CC vehicle. This sequence represents signal sequence based peptide 1, a
 CC chimeric peptide of the hydrophobic terminal domain of the viral gp41
 CC protein and the nuclear localisation signal from simian virus 40
 CC (SV40) large antigen, one of the membrane translocation peptides tested
 CC in the invention.

XX
 SQ Sequence 27 AA;
 Query Match 24.2%; Score 8; DB 23; Length 27;
 Best Local Similarity 100.0%; Pred. No. 0.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LFLGLWLG 14
 Db 3 LFLGLWLG 10

RESULT 12
 ABG75506
 ID ABG75506 standard; Peptide: 27 AA.
 AC
 XX
 XX
 DT 14-APR-2003 (first entry)
 DE Signal-sequence-based peptide 1.
 XX
 XX Translocation; delivery vehicle; agent-membrane translocation sequence;
 KW MTS; membrane translocation sequence; plasma membrane; red blood cell;
 KW immunisation; antigen; intracellular delivery; therapeutic; HIV-1;
 KW trans-activating protein; Tat; Antennapedia homeodomain protein; Antp-HD;
 KW VP22 protein; HSV-VP22; signal-sequence-based peptide; transporan; NUS;
 KW Amphiphilic model peptide; SV40; Simian virus 40; gp41;
 XX nuclear localisation sequence.

XX Chimeric - Viruses sp.
 OS Chimeric - Rhesus macaque polyoma virus
 XX
 PN US2002151004-A1.
 PD 17-OCT-2002.
 XX
 PD 16-FEB-2001; 2001US-0785802.
 PF
 XX 24-JUL-2000; 2000GB-0002848.
 PR 09-APR-2000; 2000GB-0003056.
 PR 22-DEC-2000; 2000US-0748063.
 PR 22-DEC-2000; 2000US-0748789.
 XX
 XX (CRAI/) CRAIG R.
 PA
 XX Craig R;
 PI
 XX
 XX WPI: 2003-182503/18.
 DR
 XX
 XX Preparing delivery vehicles, particularly red blood cells, for the
 PT intracellular delivery of a therapeutic agent to a target site, by
 PT loading a cell with an agent-MTS conjugate, which comprises a membrane
 PT translocation sequence
 XX
 PS Disclosure; Page 8; 43pp; English.
 XX
 XX The invention discloses a method for preparing a delivery vehicle for
 CC delivering an agent to a target site in a vertebrate. The method

CC comprises loading a cell with an agent-membrane translocation sequence
 CC (MTS) conjugate, which contains a membrane translocation sequence
 CC enabling the agent to cross the plasma membrane of a cell. Also disclosed
 CC is a pharmaceutical composition comprising a red blood cell for delivery
 CC of an agent to a vertebrate, the red blood cell comprising the novel
 CC agent-MTS conjugate and a method of immunisation of an animal with an
 CC antigen. The method is useful for preparing delivery vehicles,
 CC particularly a red blood cell, for the intracellular delivery of a
 CC therapeutic agent to a target site. The method is particularly useful for
 CC enabling an agent to cross the plasma membrane of a target cell, and for
 CC selectively releasing the agent-MTS conjugate at a target site to
 CC facilitate the uptake of the agent by the cells at the target site. The
 CC preferred MTS sequences are from HIV-1 trans-activating protein (Tat),
 CC Drosophila Antennapedia homeodomain protein (Antp-HD), Herpes Simplex-1
 CC virus VP22 protein (HSV-VP22), a signal-sequence-based peptide and a
 CC Transporan and Amphiphilic model peptide. The sequence presented is the
 CC Signal-sequence-based peptide 1, which comprises the hydrophobic terminal
 CC domain of the viral gp41 protein and the nuclear localisation sequence
 CC (NLS) from the SV40 large antigen.

XX
 SQ Sequence 27 AA;
 Query Match 24.2%; Score 8; DB 24; Length 27;
 Best Local Similarity 100.0%; Pred. No. 0.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LFLGLWLG 14
 Db 3 LFLGLWLG 10

RESULT 13
 ABG82542
 ID ABG82542 standard; peptide; 27 AA.
 XX
 AC ABG82542;
 XX
 DT 04-FEB-2003 (first entry)
 DE Signal-sequence based peptide 1 MPS peptide.
 XX
 XX Red blood cell; transgenic; cytostatic; immunosuppressive; SV40; MPS;
 KW anti-rheumatic; antiarthritic; dermatological; antithyroid; nootropic;
 KW neuroprotective; cerebroprotective; cardiac; antiarrhythmic; litholytic;
 KW antiinflammatory; nephrotropic; antidiabetic; antiasthmatic; vilerary;
 KW analgesic; osteopathic; gp41.
 XX
 XX Synthetic.
 OS Rhesus macaque polyoma virus.
 XX
 PN WO200257436-A2.
 XX
 PD 25-JUL-2002.
 XX
 XX 16-JAN-2002; 2002WO-GH00169.
 XX
 XX 19-JAN-2001; 2001GB-0001459.
 PR 23-JAN-2001; 2001US-264808P.
 PR 01-FEB-2001; 2001WO-GH00417.
 PR 16-FEB-2001; 2001US-0785802.
 PR 07-MAR-2001; 2001GB-0005631.
 PR 29-MAR-2001; 2001US-279803P.
 PR 24-JUL-2001; 2001WO-GH03327.
 XX
 XX (GENE-) GENDEL LTD.
 PA
 XX McHale AP, Craig RK;
 XX
 XX WPI: 2003-040499/03.
 DR
 XX
 XX Producing red blood cells for polypeptide delivery, by obtaining a red
 PT blood cell containing the polypeptide from a transgenic animal
 PT expressing a gene encoding the polypeptide, and sensitizing the cell

PT for disruption -

XX PS Disclosure: Page 50; 125pp; English.

XX CC The invention relates to producing a red blood cell suitable for delivery of a polypeptide to a vertebrate, by providing a transgenic animal carrying and expressing a transgene encoding the polypeptide, obtaining a red blood cell containing the polypeptide from the animal, and sensitizing the red blood cell to render it susceptible to disruption by an energy source. The polypeptide delivered using a sensitized red blood cell produced by or obtainable from the method is useful for treating diseases including cancer, non-malignant tumors, localized infections, autoimmune disorders (e.g. rheumatoid arthritis, lupus erythematosus, Graves's disease), neurodegenerative disorders (e.g. Alzheimer's disease, amyotrophic lateral sclerosis and others), neuromuscular disorders (e.g. muscular dystrophy), stroke, heart disease (heart attack or arrhythmia), arterial or venous disease (e.g. blood clots or other arterial/venous blockage, such as from cholesterol or amyloid deposition), liver disease (e.g. cirrhosis), renal disease (kidney stones, bleeding or failure), diabetes or pulmonary disease (e.g. emphysema, asthma or bronchitis). The polypeptide is also useful in the treatment of injury, inflammation, burns, carpal tunnel syndrome or conditions resulting from injury, such as osteoarthritis. The present sequence represents a MPS peptide (signal sequence derived peptide 1), a chimera of the hydrophobic terminal domain of the viral gp41 protein and the nuclear localisation signal sequence (NLS) from the SV40 large antigen. This has been found to be active in membrane translocation.

XX SQ Sequence 27 AA:

Query Match 24.2%; Score 8; DB 24; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LFLGWLGA 14
11111111

Db 3 LFLGWLGA 10

RESULT 14

ABB65236

ID ABB65236 standard; Protein: 178 AA.

XX AC ABB65236;

XX AC

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 22500.

XX Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

KW Drosophila melanogaster.

XX OS

XX WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI: 2001-656860/75.

DR N-PSDB: ABL09339.

XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.

PT

XX PS Disclosure: SEQ ID NO 22500; 21pp - Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB116176-AB130511), expressed DNA sequences (AB101940-AB116175) and the encoded proteins (ABB57737-ABB72072).

XX CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 178 AA:

Query Match 24.2%; Score 8; DB 22; Length 178;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 DRFYLGP 22
11111111

Db 108 DRFYLGP 115

RESULT 15

AAAY70761

ID AAAY70761 standard; Protein: 221 AA.

XX AC AAAY70761;

XX DT 24-JUL-2000 (first entry)

XX DE Human beta-amyloid peptide (BAP) binding protein, BBP3.

XX KW Beta-amyloid peptide binding protein; BBP; BAP; tumour suppressor; G-protein coupled receptor; GPCR; integral membrane protein; antigen; neuronal cell; nonhuman primate; NHP; G-protein signalling pathway; apoptosis; immunogen; therapeutic; treatment; prevention; diagnostic.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Domain 154..175 /label= Transmembrane_domain_1

FT Domain 176..178 /label= DRF motif

FT Domain 190..215 /note= Substitution of the Arg abrogates protection*

FT Domain /label= Transmembrane_domain_2

XX PN WC200022125-A2.

XX PJ 20-APR-2000.

XX PF 13-OCT-1999; 99WO-US21621.

XX PR 13-OCT-1998; 98US-0104104.

XX PA (AMHP) AMERICAN HOME PROD CORP.

XX PI Ozenberger BA, Kajkowski EM, Lo CF;

XX WPI: 2000-317982/27.

DR N-PSDB: AA252371.

XX Novel G-protein-coupled receptor-like proteins and polynucleotides useful for regulating apoptosis, comprises integral membrane protein traversing the membrane twice.

PT

XX Example 1; Page 67-68; 68pp; English.

XX PS

CC The present sequence is the beta-amyloid peptide (BAP) binding protein-3
CC (BBP3). It is an integral membrane protein, that traverse the membrane
CC twice. It is related to G protein-coupled receptor (GPCR) protein
CC superfamily. It interacts with G-alpha proteins and regulates the
CC activity of G-protein signalling pathways. BBP genes are widely expressed
CC in neuronal cells of nonhuman primate (NHP) brain and overexpressed in
CC some tumours. It functions as a suppressor of apoptosis induction. BBP
CC proteins are used as immunogens to raise antibodies, useful as
CC therapeutics and as antigens in solid phase assays. They are also useful
CC as reagents to identify molecules which effect the interaction of BBP and
CC a cloned protein, that are useful in the treatment or prevention of
CC diseases associated with apoptosis. The polynucleotides are useful for
CC diagnostics.

xx

SQ Sequence 221 AA:

Query Match 24.2%; Score 8; DB 21; Length 221;
Best Local Similarity 100.0%; Pred. No. 5.7;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 GADREYLG 20

|||||||

Db 174 GADREYLG 181

Search completed: September 26, 2003, 17:39:25

Job time : 17.7106 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:36:55 : Search time 5.89787 Seconds
(without alignments)
236.739 Million cell updates/sec

Title: US-09-852-100A-2_COPY_185_217

Perfect score: 33

Sequence: 1 VAVLSLFLGW:GADREYLGYPALGLLK:CTVG 33

Scoring table: OUIGO

Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents_MA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/iaa/6R_COMB.pep:*
- 5: /cgn2_6/ptodata/1/iaa/PCBUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	24.2	358	4	US-09-411-628-11
2	7	21.2	451	4	US-09-345-473E-45
3	7	21.2	462	4	US-09-328-352-6889
4	7	21.2	483	4	US-09-252-991A-29267
5	7	21.2	617	4	US-09-345-473E-44
6	6	18.2	46	4	US-09-489-847-131
7	6	18.2	70	4	US-09-205-258-293
8	6	18.2	72	3	US-09-905-223-314
9	6	18.2	101	4	US-09-489-847-306
10	6	18.2	117	4	US-09-228-986-111
11	6	18.2	123	3	US-08-675-508-2
12	6	18.2	123	3	US-09-203-939-2
13	6	18.2	123	3	US-09-203-939-6
14	6	18.2	123	3	US-09-251-835-2
15	6	18.2	123	3	US-09-251-835-6
16	6	18.2	123	3	US-09-318-503-2
17	6	18.2	123	3	US-09-318-503-6
18	6	18.2	123	3	US-09-038-261A-2
19	6	18.2	123	3	US-09-038-261A-6
20	6	18.2	123	4	US-09-564-329A-2
21	6	18.2	123	4	US-09-564-329A-6
22	6	18.2	209	2	US-08-808-550-31
23	6	18.2	261	4	US-09-252-991A-17349
24	6	18.2	281	4	US-09-252-991A-33042
25	6	18.2	325	4	US-09-252-991A-16657
26	6	18.2	349	4	US-09-162-524-3
27	6	18.2	350	4	US-08-637-670-25

28 6 18.2 360 4 US-09-328-352-5126 Sequence 5126, Ap
29 6 18.2 380 4 US-09-252-991A-24774 Sequence 24774, A
30 6 18.2 384 1 US-08-221-750A-11 Sequence 11, Appl
31 6 18.2 389 4 US-09-252-991A-21202 Sequence 21202, A
32 6 18.2 396 4 US-09-198-452A-147 Sequence 147, Ap
33 6 18.2 404 4 US-09-328-352-6747 Sequence 6747, Ap
34 6 18.2 421 4 US-09-252-991A-30177 Sequence 30177, A
35 6 18.2 430 4 US-09-252-991A-31388 Sequence 31388, A
36 6 18.2 432 1 US-08-476-008-61 Sequence 61, Appl
37 6 18.2 432 1 US-08-306-363-61 Sequence 61, Appl
38 6 18.2 432 1 US-08-833-485-61 Sequence 61, Appl
39 6 18.2 432 3 US-09-137-440-61 Sequence 61, Appl
40 6 18.2 434 1 US-08-162-392-4 Sequence 4, Appl
41 6 18.2 434 1 US-08-643-300-4 Sequence 4, Appl
42 6 18.2 434 1 US-08-643-297-4 Sequence 4, Appl
43 6 18.2 434 1 US-08-643-298-4 Sequence 4, Appl
44 6 18.2 434 2 US-08-643-301-4 Sequence 4, Appl
45 6 18.2 434 5 PCT-US94-14095-4 Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-411-628-11
: Sequence 11, Application US/09411628
: Patent No. 6428994
: GENERAL INFORMATION:
: APPLICANT: University of Southern California
: TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
: FILE REFERENCE: 13761-707
: CURRENT APPLICATION NUMBER: US/09/411,628
: CURRENT FILING DATE: 1999-10-01
: EARLIER APPLICATION NUMBER: US 60/102,906
: EARLIER FILING DATE: 1998-10-02
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 11
: LENGTH: 358
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-411-628-11

Query Match: 24.2% Score 8; DB 4; Length 358;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative C; Mismatches 0; Indels 0; Gaps 0;

Qy 21 YPALGLLK 28
Pb 259 YPALGLLK 265

RESULT 2
US-09-345-473E-45
: Sequence 45, Application US/09345473E
: Patent No. 6558903
: GENERAL INFORMATION:
: APPLICANT: Hodge, Martin
: TITLE OF INVENTION: No. 6558903el Kinases and Uses Thereof
: FILE REFERENCE: 35800/183781
: CURRENT APPLICATION NUMBER: US/09/345,473E
: CURRENT FILING DATE: 1999-06-30
: NUMBER OF SEQ ID NOS: 62
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 45
: LENGTH: 451
: TYPE: PRT
: ORGANISM: Mus musculus
US-09-345-473E-45

Query Match: 21.2% Score 7; DB 4; Length 451;
Best Local Similarity 100.0%; Pred. No. 21;


```

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 ALSFLG 10
Db 416 ALSFLG 422

RESULT 3
US-09-328-352-6888
: Sequence 6888, Application US/09328352
: Patent No. 6562958
: GENERAL INFORMATION:
: APPLICANT: Gary L. Berton et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
: FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
: CURRENT APPLICATION NUMBER: US/09-03PA
: CURRENT FILING DATE: 1999-06-04
: NUMBER OF SEQ ID NOS: 352
: SEQ ID NO 588
: LENGTH: 462
: TYPE: PRT
: ORGANISM: Acinetobacter baumannii
US-09-328-352-6888

Query Match 21.2%; Score 7; DB 4; Length 462;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 ALSFLG 10
Db 334 ALSFLG 340

RESULT 4
US-09-252-991A-29267
: Sequence 29267, Application US/09252991A
: Patent No. 6551795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: CURRENT APPLICATION NUMBER: US/09-02-18
: CURRENT FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 29267
: LENGTH: 483
: TYPE: PRT
: ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29267

Query Match 21.2%; Score 7; DB 4; Length 483;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 22 PALGLLK 28
Db 438 PALGLLK 444

RESULT 5
US-09-345-473E-44
: Sequence 44, Application US/09345473E
: Patent No. 6558903
: GENERAL INFORMATION:
: APPLICANT: Hodge, Martin
: TITLE OF INVENTION: No. 6558903ei Kinases and Uses Thereof
: FILE REFERENCE: 35800/183781

```

```

: CURRENT APPLICATION NUMBER: US/09/345,473E
: CURRENT FILING DATE: 1999-06-30
: NUMBER OF SEQ ID NOS: 62
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 44
: LENGTH: 617
: TYPE: PRT
: ORGANISM: Mus musculus
US-09-345-473E-44

Query Match 21.2%; Score 7; DB 4; Length 617;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 ALSFLG 10
Db 582 ALSFLG 588

RESULT 6
US-09-489-847-131
: Sequence 131, Application US/09489847
: Patent No. 6475195
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: 98 Human Secreted Proteins
: FILE REFERENCE: P2031E1
: CURRENT APPLICATION NUMBER: US/09/489,847
: CURRENT FILING DATE: 2000-01-24
: EARLIER APPLICATION NUMBER: PCT/US99/17130
: EARLIER FILING DATE: 1999-07-29
: EARLIER APPLICATION NUMBER: 60/094,657
: EARLIER FILING DATE: 1998-07-30
: EARLIER APPLICATION NUMBER: 60/095,486
: EARLIER FILING DATE: 1998-08-05
: EARLIER APPLICATION NUMBER: 60/096,319
: EARLIER FILING DATE: 1998-08-12
: EARLIER APPLICATION NUMBER: 60/095,454
: EARLIER FILING DATE: 1998-08-06
: EARLIER APPLICATION NUMBER: 60/095,455
: EARLIER FILING DATE: 1998-08-06
: NUMBER OF SEQ ID NOS: 376
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 131
: LENGTH: 46
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (46)
: OTHER INFORMATION: Xaa equals stop translation
US-09-489-847-131

Query Match 18.2%; Score 6; DB 4; Length 46;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 24 LGLLKF 29
Db 10 LGLLKF 15

RESULT 7
US-09-205-258-298
: Sequence 298, Application US/09205258
: Patent No. 6525174
: GENERAL INFORMATION:
: APPLICANT: Young et al.
: TITLE OF INVENTION: 207 Human Secreted Proteins
: FILE REFERENCE: P2007P1
: CURRENT APPLICATION NUMBER: US/09/205,258
: CURRENT FILING DATE: 1998-12-04
: EARLIER APPLICATION NUMBER: PCT/US98/11422

```

```
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
```

```
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 298
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (70)
; OTHER INFORMATION: Xaa equals stop translation
US-09-205-258-298

Query Match      18.2%; Score 6; DB 4; Length 70;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY      4 ALSLFL 9
      |||||
CB      4 ALSLFL 9

RESULT 8
US-08-305-223-314
; Sequence 314, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 314:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 72 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: -42...-1
; IDENTIFICATION METHOD: Von Heijne matrix
```

OTHER INFORMATION: score 6.1
OTHER INFORMATION: seq DLAVALSLPAAWT/ES
US-08-905-223-314

Query Match 18.2%, Score 6; DB 3; Length 72;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AVALSL 7
DB 32 AVALSL 37

RESULT 9

US-09-489-847-306
; Sequence 306, Application US/09499847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/1713C
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 50/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 378
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 306
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-489-847-306

Query Match 18.2%, Score 6; DB 4; Length 101;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AVALSL 7
DB 89 AVALSL 94

RESULT 10

US-09-228-986-111
; Sequence 111, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niele
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 110C0/102C
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 111
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-228-986-111

Query Match 18.2%, Score 6; DB 4; Length 117;

Best Local Similarity 100.0%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAVALS 6
DB 37 VAVALS 42

RESULT 11

US-06-675-508-2
; Sequence 2, Application US/08675508
; Patent No. 5856136
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: NOVEL HUMAN STEM CELL ANTIGENS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,508
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0066 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 123 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: SCAN-2
; CLONE:
US-08-675-508-2

Query Match 18.2%, Score 6; DB 2; Length 123;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 PALGLL 27
DB 111 PALGLL 116

RESULT 12

US-09-203-939-2
; Sequence 2, Application US/09203939
; Patent No. 6258939
; GENERAL INFORMATION:
; APPLICANT: Reiter, Robert E.
; APPLICANT: Witte, Owen N.
; TITLE OF INVENTION: PSMA: PROSTATE STEM CELL ANTIGEN AND USES THEREOF
; FILE REFERENCE: 30435.54US11
; CURRENT APPLICATION NUMBER: US/09/203,939
; CURRENT FILING DATE: 2000-12-02
; PRIOR APPLICATION NUMBER: 08/814,279
; PRIOR FILING DATE: 1997-03-10
; PRIOR APPLICATION NUMBER: 60/071,141

```

; PRIOR FILING DATE: 1998-01-12
; PRIOR APPLICATION NUMBER: 60/074,675
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 09/038,261
; PRIOR FILING DATE: 1998-03-10
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 123
; TYPE: PRT
; ORGANISM: HUMAN PSCA (hPSCA)
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (50)..(64)
; NAME/KEY: SITE
; LOCATION: (71)..(82)
; NAME/KEY: SITE
; LOCATION: (67)..(81)
US-09-203-939-2

```

```

Query Match      18.2%  Score 6;  DB 3;  Length 123;
Best Local Similarity 100.0%;  Pred. No. 63;
Matches 6;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

QY      22 PALGGL 27
DB      111 PALGGL 116

```

RESULT 13

```
US-09-203-939-5
```

```

; Sequence 6, Application US/09201939
; Patent No. 6258939
; GENERAL INFORMATION:
; APPLICANT: Reiter, Robert E.
; APPLICANT: Witte, Owen N.
; TITLE OF INVENTION: PSCA: PROSTATE STEM CELL ANTIGEN AND USES THEREOF
; FILE REFERENCE: 30435.54US1
; CURRENT APPLICATION NUMBER: US/09/203,939
; PRIOR FILING DATE: 2000-12-02
; PRIOR APPLICATION NUMBER: 08/814,279
; PRIOR FILING DATE: 1997-03-10
; PRIOR APPLICATION NUMBER: 60/071,141
; PRIOR FILING DATE: 1998-01-12
; PRIOR APPLICATION NUMBER: 60/074,675
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 09/038,261
; PRIOR FILING DATE: 1998-03-10
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 123
; TYPE: PRT
; ORGANISM: HUMAN PSCA (hPSCA)
US-09-203-939-6

```

```

Query Match      18.2%  Score 6;  DB 3;  Length 123;
Best Local Similarity 100.0%;  Pred. No. 63;
Matches 6;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

QY      22 PALGGL 27
DB      111 PALGGL 116

```

RESULT 14

```
US-09-251-835-2
```

```

; Sequence 2, Application US/09251835A
; Patent No. 6261789
; GENERAL INFORMATION:
; APPLICANT: Reiter, Robert E.
; APPLICANT: Witte, Owen N.
; TITLE OF INVENTION: PSCA: PROSTATE STEM CELL ANTIGEN

```

```

; FILE REFERENCE: 30435.54US12
; CURRENT APPLICATION NUMBER: US/09/251,835A
; CURRENT FILING DATE: 1999-02-17
; PRIOR APPLICATION NUMBER: 08/814,279
; PRIOR FILING DATE: 1997-03-10
; PRIOR APPLICATION NUMBER: 60/071,141
; PRIOR FILING DATE: 1998-01-12
; PRIOR APPLICATION NUMBER: 60/074,675
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 09/038,261
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 09/203,939
; PRIOR FILING DATE: 1998-12-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 123
; TYPE: PRT
; ORGANISM: HUMAN PSCA (hPSCA)
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (50)..(64)
; NAME/KEY: SITE
; LOCATION: (71)..(82)
; NAME/KEY: SITE
; LOCATION: (67)..(81)
US-09-251-835-2

```

```

Query Match      18.2%  Score 6;  DB 3;  Length 123;
Best Local Similarity 100.0%;  Pred. No. 63;
Matches 6;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

QY      22 PALGGL 27
DB      111 PALGGL 116

```

RESULT 15

```
US-09-251-835-6
```

```

; Sequence 6, Application US/09251835A
; Patent No. 6261789
; GENERAL INFORMATION:
; APPLICANT: Reiter, Robert E.
; APPLICANT: Witte, Owen N.
; TITLE OF INVENTION: PSCA: PROSTATE STEM CELL ANTIGEN
; FILE REFERENCE: 30435.54US12
; CURRENT APPLICATION NUMBER: US/09/251,835A
; CURRENT FILING DATE: 1999-02-17
; PRIOR APPLICATION NUMBER: 08/814,279
; PRIOR FILING DATE: 1997-03-10
; PRIOR APPLICATION NUMBER: 60/071,141
; PRIOR FILING DATE: 1998-01-12
; PRIOR APPLICATION NUMBER: 60/074,675
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 09/038,261
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 09/203,939
; PRIOR FILING DATE: 1998-12-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 123
; TYPE: PRT
; ORGANISM: HUMAN PSCA (hPSCA)
US-09-251-835-6

```

```

Query Match      18.2%  Score 6;  DB 3;  Length 123;
Best Local Similarity 100.0%;  Pred. No. 63;
Matches 6;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

QY      22 PALGGL 27
DB      111 PALGGL 116

```

Search completed: September 26, 2003, 17:43:24
Job time : 6.89787 secs



GenCore version 5.1.6
Copyright. (c) 1993 - 2003 Computer, Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:41:48 : Search time 11.5149 seconds
(without alignments)
433.624 Million cell updates/sec

Title: US-09-852-100A-2_COPY_185_217

Perfect score: 33

Sequence: 1 VAVALSLFLGWLGLADRFYLGYPALGLLKFCITVG 33

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 566894 seqs, 151307093 residues

Word size : 0

Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PublishedApplications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	100.0	269	9	US-09-852-100A-2
2	33	100.0	269	10	US-09-833-503A-2
3	33	100.0	269	15	US-10-199-881-2
4	8	24.2	21	15	US-10-226-956-288
5	8	24.2	21	15	US-10-211-088-306
6	8	24.2	24	11	US-09-015-914B-28
7	8	24.2	27	10	US-09-785-802A-8
8	8	24.2	27	11	US-09-965-876A-2
9	8	24.2	27	11	US-09-915-914B-27
10	8	24.2	27	15	US-10-077-555-8
11	8	24.2	221	11	US-09-833-503A-6
12	8	24.2	221	11	US-09-992-600A-82
13	8	24.2	221	11	US-09-924-340-82
14	8	24.2	221	11	US-09-974-879-230
15	8	24.2	221	12	US-09-992-095B-82

16	8	24.2	221	12	US-10-154-678-32	Sequence 82, Appl
17	8	24.2	221	12	US-09-999-570-82	Sequence 82, Appl
18	8	24.2	221	15	US-10-000-489-82	Sequence 82, Appl
19	8	24.2	221	15	US-10-000-986-82	Sequence 82, Appl
20	8	24.2	221	15	US-10-199-881-6	Sequence 6, Appl
21	8	24.2	221	11	US-09-305-736-230	Sequence 230, Appl
22	8	24.2	230	11	US-09-796-753-50	Sequence 50, Appl
23	8	24.2	247	11	US-09-796-753-48	Sequence 48, Appl
24	8	24.2	358	12	US-10-174-794-11	Sequence 11, Appl
25	7	21.2	122	9	US-09-867-550-80	Sequence 80, Appl
26	7	21.2	173	11	US-09-946-374-332	Sequence 332, Appl
27	7	21.2	173	12	US-10-015-387A-332	Sequence 332, Appl
28	7	21.2	173	12	US-10-006-130A-332	Sequence 332, Appl
29	7	21.2	173	12	US-10-006-172A-332	Sequence 332, Appl
30	7	21.2	173	12	US-10-015-392A-332	Sequence 332, Appl
31	7	21.2	173	12	US-10-017-253A-332	Sequence 332, Appl
32	7	21.2	173	12	US-10-017-306A-332	Sequence 332, Appl
33	7	21.2	173	15	US-10-006-856A-332	Sequence 332, Appl
34	7	21.2	173	15	US-10-006-818A-332	Sequence 332, Appl
35	7	21.2	173	15	US-10-015-393A-332	Sequence 332, Appl
36	7	21.2	173	15	US-10-015-869A-332	Sequence 332, Appl
37	7	21.2	173	15	US-10-012-121A-332	Sequence 332, Appl
38	7	21.2	173	15	US-10-006-116A-332	Sequence 332, Appl
39	7	21.2	173	15	US-10-006-117A-332	Sequence 332, Appl
40	7	21.2	173	15	US-10-017-527A-332	Sequence 332, Appl
41	7	21.2	173	15	US-10-013-913A-332	Sequence 332, Appl
42	7	21.2	173	15	US-10-007-194A-332	Sequence 332, Appl
43	7	21.2	173	15	US-10-013-430A-332	Sequence 332, Appl
44	7	21.2	173	15	US-10-011-671A-332	Sequence 332, Appl
45	7	21.2	173	15	US-10-012-755A-332	Sequence 332, Appl

ALIGNMENTS

RESULT :
US-09-852-100A-2
: Sequence 2, Application US/09852100A
: Patent No. US20020058267A1
: GENERAL INFORMATION:
: APPLICANT: American Home Products
: TITLE OF INVENTION: Beta-amyloid Peptide-Binding Proteins and Polynucleotides Enc
: FILE OF INVENTION: Same
: FILE REFERENCE: AHP981261P2
: CURRENT APPLICATION NUMBER: US/09/852,100A
: CURRENT FILING DATE: 2001-05-09
: PRIOR APPLICATION NUMBER: US 09/172,990
: PRIOR FILING DATE: 1998-10-14
: PRIOR APPLICATION NUMBER: US 60/104,104
: PRIOR FILING DATE: 1998-10-13
: PRIOR APPLICATION NUMBER: PTC/US99/21621
: PRIOR FILING DATE: 1999-10-13
: PRIOR APPLICATION NUMBER: US 09/060,609
: PRIOR FILING DATE: 1998-04-15
: PRIOR APPLICATION NUMBER: US 60/054,583
: PRIOR FILING DATE: 1997-04-16
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 2
: LENGTH: 269
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-852-100A-2

Query Match 100.0%; Score 33; DB 9; Length 269;
Best Local Similarity 100.0%; Pred. No. 9.4e-24;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAVALSLFLGWLGLADRFYLGYPALGLLKFCITVG 33
|||||
DB 185 VAVALSLFLGWLGLADRFYLGYPALGLLKFCITVG 217

```
RESULT 2
US-09-833-503A-2
; Sequence 2, Application US/09833503A
; Patent No. US20020146760A1
; GENERAL INFORMATION:
; APPLICANT: Ozenberger, Bradley A
; APPLICANT: Kajkowski, Eileen M
; APPLICANT: Lo, Ching-Hsiung F
; APPLICANT: American Home Products Corporation
; TITLE OF INVENTION: No. US20020146760A1el G-Protein-Coupled Receptor-Like Proteins and Methods of Using
; TITLE OF INVENTION: Polynucleotides Encoded By Them, and Methods of Using
; TITLE OF INVENTION: Same
; FILE REFERENCE: AHP98155-00PCT
; CURRENT APPLICATION NUMBER: US/09/833,503A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/104,104
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-503A-2

Query Match      100.0%; Score 33; DB 10; Length 269;
Best Local Similarity 100.0%; Pred. No. 9,4e-24;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAVALSFLGWLGNADRYLGYPALGLLKFTCTVG 33
      |||
DB 185 VAVALSFLGWLGNADRYLGYPALGLLKFTCTVG 217

RESULT 3
US-10-199-881-2
; Sequence 2, Application US/10199881
; Publication No. US20030096356A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: No. US20030096356A1el G-Protein-Coupled Receptor-Like Proteins and Methods of Using Same
; FILE REFERENCE: AHP98155C1
; CURRENT APPLICATION NUMBER: US/10/199,881
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/US95/21621
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: US 90/833,5081
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 60/104,104
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-881-2

Query Match      100.0%; Score 33; DB 15; Length 269;
Best Local Similarity 100.0%; Pred. No. 9,4e-24;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAVALSFLGWLGNADRYLGYPALGLLKFTCTVG 33
      |||
DB 185 VAVALSFLGWLGNADRYLGYPALGLLKFTCTVG 217

RESULT 4
US-10-226-956-288
; Sequence 288, Application US/10226955
; Publication No. US20030060399A1
; GENERAL INFORMATION:
; APPLICANT: Divita, Gilles
; APPLICANT: Fernandez, Joseph
; APPLICANT: Heitz, Frederic
```

```
; APPLICANT: Brophy, Colleen
; APPLICANT: Komalavilas, Padmini
; APPLICANT: Panitch, Alyssa
; APPLICANT: Joshi, Lokesh
; APPLICANT: Seal, Brandon L.
; TITLE OF INVENTION: REAGENTS AND METHODS FOR SMOOTH MUSCLE THERAPIES
; FILE REFERENCE: ASU-106-US
; CURRENT APPLICATION NUMBER: US/10/226,956
; CURRENT FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: 60/314,535
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 320
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 288
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-226-956-288

Query Match      24.2%; Score 8; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LFLGWLGA 14
      |||
DB 3 LFLGWLGA 10

RESULT 5
US-10-211-088-306
; Sequence 306, Application US/10211088
; Publication No. US20030104479A1
; GENERAL INFORMATION:
; APPLICANT: Bright, Gary R.
; APPLICANT: Premkumar, D. David
; APPLICANT: Chen, Yih-Tai
; TITLE OF INVENTION: No. US20030104479A1el Fusion Proteins And Assays For Molecular
; FILE REFERENCE: 01-1022-US
; CURRENT APPLICATION NUMBER: US/10/211,088
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/309,395
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/341,589
; PRIOR FILING DATE: 2001-12-13
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 306
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Protein-derived transport peptide
US-10-211-088-306

Query Match      24.2%; Score 8; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LFLGWLGA 14
      |||
DB 3 LFLGWLGA 10

RESULT 6
US-09-915-914B-28
; Sequence 28, Application US/09915914B
; Publication No. US20030119725A1
; GENERAL INFORMATION:
; APPLICANT: Divita, Gilles
; APPLICANT: Fernandez, Joseph
; APPLICANT: Heitz, Frederic
```

```
; APPLICANT: Morris, May
; APPLICANT: Mery, Jean
; APPLICANT: Archdeacon, John
; APPLICANT: Horndorp, Kyle
; TITLE OF INVENTION: PEPTIDE-MEDIATED DELIVERY OF MOLECULES INTO CELLS
; FILE REFERENCE: AM-00105.P.1.1
; CURRENT APPLICATION NUMBER: US/09/915,914B
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: US 60/221,932
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic sequence
US-09-915-914B-28

Query Match      24.2%; Score 8; DB 11; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.63;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      7 LFLGWLGA 14
Db      3 LFLGWLGA 10
|||||||

RESULT 7
US-09-785-802A-8
; Sequence 8, Application US/09785802A
; Patent No. US20020151004A1
; GENERAL INFORMATION:
; APPLICANT: Craig, Roger
; TITLE OF INVENTION: DELIVERY VEHICLES AND METHODS FOR USING THE SAME
; FILE REFERENCE: 11067/2035
; CURRENT APPLICATION NUMBER: US/09/785,802A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 09/748,06
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/748,789
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-785-802A-8

Query Match      24.2%; Score 8; DB 10; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.63;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      7 LFLGWLGA 14
Db      3 LFLGWLGA 10
|||||||

RESULT 8
US-09-965-876A-2
; Sequence 2, Application US/09965876A
; Publication No. US20030096243A1
; GENERAL INFORMATION:
; APPLICANT: Cellomics, Inc.
; APPLICANT: Busa, William B
; TITLE OF INVENTION: Methods and Reagents for Live-cell Gene Expression Quantification
; FILE REFERENCE: 00-789-A
; CURRENT APPLICATION NUMBER: US/09/965,876A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/236,407
; PRIOR FILING DATE: 2000-09-28

; APPLICANT: Divita, Gilles
; APPLICANT: Fernandez, Joseph
; APPLICANT: Heitz, Frederic
; APPLICANT: Morris, May
; APPLICANT: Mery, Jean
; APPLICANT: Archdeacon, John
; APPLICANT: Horndorp, Kyle
; TITLE OF INVENTION: PEPTIDE-MEDIATED DELIVERY OF MOLECULES INTO CELLS
; FILE REFERENCE: AM-00105.P.1.1
; CURRENT APPLICATION NUMBER: US/09/915,914B
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: US 60/221,932
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic sequence
US-09-915-914B-27

Query Match      24.2%; Score 8; DB 11; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.63;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      7 LFLGWLGA 14
Db      3 LFLGWLGA 10
|||||||

RESULT 9
US-09-915-914B-27
; Sequence 27, Application US/09915914B
; Publication No. US20030119725A1
; GENERAL INFORMATION:
; APPLICANT: Divita, Gilles
; APPLICANT: Fernandez, Joseph
; APPLICANT: Heitz, Frederic
; APPLICANT: Morris, May
; APPLICANT: Mery, Jean
; APPLICANT: Archdeacon, John
; APPLICANT: Horndorp, Kyle
; TITLE OF INVENTION: PEPTIDE-MEDIATED DELIVERY OF MOLECULES INTO CELLS
; FILE REFERENCE: AM-00105.P.1.1
; CURRENT APPLICATION NUMBER: US/09/915,914B
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: US 60/221,932
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic sequence
US-09-915-914B-27

Query Match      24.2%; Score 8; DB 11; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.63;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      7 LFLGWLGA 14
Db      3 LFLGWLGA 10
|||||||

RESULT 10
US-10-077-555-8
; Sequence 8, Application US/10077555
; Publication No. US20030077289A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Rong-fu
; TITLE OF INVENTION: Use of Cell-Penetrating Peptides to Generate Antitumor Immun;
; FILE REFERENCE: P02373US1/10200806
; CURRENT APPLICATION NUMBER: US/10/077,555
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/268,687
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 27
; TYPE: PRT
```



```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-077-555-8

Query Match      24.2%  Score 8;  DB 15;  Length 27;
Best Local Similarity 100.0%;  Pred. No. 0; 53;
Matches      8;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Oy  7  LFLGWLGA 14
    |||||||
Db   3  LFLGWLGA 10

RESULT 11
US-09-833-503A-6
; Sequence 6, Application US/09833503A
; Patent No. US20020146760A1
; GENERAL INFORMATION:
; APPLICANT: Ozenberger, Bradley A
; APPLICANT: Kalkowski, Eileen M
; APPLICANT: Lo, Ching-Hsiung F
; APPLICANT: American Home Products Corporation
; TITLE OF INVENTION: No. US20020146760A1e1 G-protein-Coupled Receptor-like Proteins and
; TITLE OF INVENTION: Polynucleotides Encoded By Them, and Methods of Using
; TITLE OF INVENTION: Same
; FILE REFERENCE: AHP98165-00PCT
; CURRENT APPLICATION NUMBER: US/09/833,503A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/104,104
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-503A-6

Query Match      24.2%  Score 8;  DB 10;  Length 221;
Best Local Similarity 100.0%;  Pred. No. 4;
Matches      8;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Oy  13  GADRFFYLG 20
    |||||||
Db   174  GADRFFYLG 18;

RESULT 12
US-09-922-600A-82
; Sequence 82, Application US/09922600A
; Publication No. US20030027161A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephanie
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.USA DIV
; CURRENT APPLICATION NUMBER: US/09/922,600A
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 63/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/51715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: JPatent

US-09-922-600A-82
; Sequence 82, Application US/09922600A
; Publication No. US20030027161A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephanie
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.USA DIV
; CURRENT APPLICATION NUMBER: US/09/922,600A
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 63/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/51715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: JPatent

US-09-922-600A-82
; Sequence 82, Application US/09922600A
; Publication No. US20030027161A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephanie
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.USA DIV
; CURRENT APPLICATION NUMBER: US/09/922,600A
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 63/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/51715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: JPatent

US-09-924-340-82
; Sequence 82, Application US/09924340
; Publication No. US20030027248A1
; GENERAL INFORMATION:
; APPLICANT: Bedjanin, Stephanie
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.USA DIV
; CURRENT APPLICATION NUMBER: US/09/924,340
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 82
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SIGNAL
; LOCATION: 1..32
US-09-924-340-82

Query Match      24.2%  Score 8;  DB 11;  Length 221;
Best Local Similarity 100.0%;  Pred. No. 4;
Matches      8;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Oy  13  GADRFFYLG 20
    |||||||
Db   174  GADRFFYLG 181

RESULT 13
US-09-924-340-82
; Sequence 82, Application US/09924340
; Publication No. US20030027248A1
; GENERAL INFORMATION:
; APPLICANT: Bedjanin, Stephanie
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.USA DIV
; CURRENT APPLICATION NUMBER: US/09/924,340
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 82
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SIGNAL
; LOCATION: 1..32
US-09-924-340-82

Query Match      24.2%  Score 8;  DB 11;  Length 221;
Best Local Similarity 100.0%;  Pred. No. 4;
Matches      8;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Oy  13  GADRFFYLG 20
    |||||||
Db   174  GADRFFYLG 181

RESULT 14
US-09-974-679-230
; Sequence 230, Application US/09974879
; Publication No. US20030028003A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 125 Human Secreted Proteins
; FILE REFERENCE: P2020P2
; CURRENT APPLICATION NUMBER: US/09/974,879
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/239,893
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 09/818,683
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 09/305,736
; PRIOR FILING DATE: 1999-05-05
```

```

; PRIOR APPLICATION NUMBER: PCT/US98/234355
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: US 60/064,911
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,912
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,983
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,900
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,988
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,987
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,908
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,984
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,985
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/066,094
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,100
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,089
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,095
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,090
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 611
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 230
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (184)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-974-879-230

```

```

Query Match      24.2%; Score 8; DB 11; Length 221;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      13 GADRFYLG 20
Db      174 GADRFYLG 181

```

```

RESULT 15
US-09-992-095B-82
; Sequence 82, Application US/0992095B
; Publication No. US20030157485A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US5.DIV
; CURRENT APPLICATION NUMBER: US/09/992.095B
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574

```

```

; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 82
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: 1..32
US-09-992-095B-82

```

```

Query Match      24.2%; Score 8; DB 12; Length 221;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      13 GADRFYLG 20
Db      174 GADRFYLG 181

```

```

Search completed: September 26, 2003, 17:55:13
Job time : 12.5149 secs

```



```

Query Match      24.2%; Score 8; DB 1: Length 358;
Best Local Similarity 100.0%; Pred. No. 1;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 YPALGSLK 28
DB 259 YPALGSLK 266

RESULT 2
ID YPJC_BACSCU STANDARD; PSZ: 215 AA.
AC P42978;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein ypjC.
GN YPJC OR J04C.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / Marburg;
RX MEDLINE=96349105; PubMed=8760912;
RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
RA Serror P.;
RA *Sequence analysis of the Bacillus subtilis chromosome region between
RT the serA and kds loci cloned in a yeast artificial chromosome.*;
RL Microbiology 142:2005-2016(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Ailoni G.,
RA Azevedo V., Bertero M.G., Bossieres P., Bilotin A., Borchert S.,
RA Borris R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.I.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N.,
RA Gham S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Katamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.W., Levine A., Liu H., Masuda S., Maestri D., Nakai S., Noback M.,
RA Medina N., Mellado R.P., Mizuno M., Moesti D., Ogawa K., Oudega B., Park S.H.,
RA Noone D., O'Reilly M., Ogawa K., Ogawa K., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Taccioni A., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka I., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzengger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RA *The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.*;
RL Nature 390:249-256(1997).
RC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: TO B.SUBTILIS YPJC AND YOFU.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; AJ250712; CAB60616.1; -

```

```

modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DB EMBL; L38424; AAA92872.1; -
DR EMBL; L47709; AAB38440.1; -
DR EMBL; Z99135; CAB14167.1; -
DR PIR; C69937; C69937;
DR S:tblist; BG11209; ypjC.
DR InterPro; IPR003740; DUF161.
DR Pfam; PF02588; DUF161; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 31 51 POTENTIAL.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 104 124 POTENTIAL.
SQ SEQUENCE 215 AA; 23582 MW; D314CF7225F8A983 CRC64;

Query Match      21.2%; Score 7; DB 1: Length 215;
Best Local Similarity 100.0%; Pred. No. 7;
Matches      7; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 3 VALSLEL 9
DB 10 VALSLEL 16

RESULT 3
ID CLDH_HUMAN STANDARD; PRZ: 224 AA.
AC P56750;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Claudin-17.
GN CLDN17.
CS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Keep T.J., Inglehearn C.F.;
RA Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Toki Y., Choi D.-K., Groner Y.,
RA Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
RA Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA Weinmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Leinhardt R., Reinhardt R., Yaspo M.-L.;
RA *The DNA sequence of human chromosome 21.*;
RL Nature 405:311-319(2000).
RC -!- FUNCTION: COMPONENT OF TIGHT JUNCTION (TJ) STRANDS.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the claudin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; AJ250712; CAB60616.1; -

```

DR EXBL; AP001707; BAA55566.1; ..
 DR Genew; HGNC:2038; CLDN17.
 DR InterPro; IPR006187; Claudin.
 DR InterPro; IPR006188; Claudin sup.
 DR InterPro; IPR004031; EMP22_Claudin.
 DR Pfam; PF00822; EMP22_Claudin; 1.
 DR PRINTS; PR01077; CLAUDIN.
 DR PROSITE; PS01346; CLAUDIN; 1.
 KW Tight junction; Transmembrane.
 FT TRANSMEM 8 28 POTENTIAL.
 FT TRANSMEM 82 102 POTENTIAL.
 FT TRANSMEM 125 145 POTENTIAL.
 FT TRANSMEM 165 185 POTENTIAL.
 FT TRANSMEM 224 AA; 24603 MW; 183JED3178B7F63A CRC64;
 SQ SEQUENCE 224 AA; 24603 MW; 183JED3178B7F63A CRC64;
 Query Match 21.2%; Score 7; DB 1; Length 224;
 Best Local Similarity 100.0%; Pred. No. 7.2;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VAVALS 7
 Db 86 VAVALS 92
 RESULT 4
 CYB_RHOVI STANDARD; PRT: 419 AA.
 ID CYB_RHOVI
 AC P81378;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DF Cytochrome b.
 GN PEB3 OR PBCB.
 OS Rhodopsinodomonas viridis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Hyphomicrobiaceae; Blastochloris.
 OX NCBI_TaxID=1079;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SRAIN=DSK133;
 RA MEDLINE=90158506; PubMed=2560136;
 RT "Cloning and sequencing of the fcb, B and C genes encoding the
 cytochrome b/c1 complex from Rhodospseudomonas viridis";
 RL Mol. Gen. Genet. 219:445-452(1989)
 CC -i- FUNCTION: Component of the ubiquinol:cytochrome c reductase
 complex (complex III or cytochrome b-c1 complex), which is a
 respiratory chain that generates an electrochemical potential
 coupled to ATP synthesis (By similarity).
 CC -i- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or B1 or
 b562) is low-potential and absorbs at about 552, and heme 2 (or B2
 or b566) is high-potential and absorbs at about 566 (By
 similarity).
 CC -i- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b.
 cytochrome c1 and the Rieske protein (By similarity).
 CC -i- SIMILARITY: Belongs to the cytochrome b family.
 PR; JC0346; J00346.
 DR InterPro; IPR005798; Cytb_b6_C.
 DR InterPro; IPR005797; Cytb_b6_N.
 DR Pfam; PF00032; cytochrome_b_c1.
 DR Pfam; PF00033; cytochrome_b_N; 1.
 DR PROSITE; PS00192; CYTOCHROME_B; 1.
 DR PROSITE; PS00193; CYTOCHROME_B; 00; 1.
 KW Electron transport; Respiratory chain; Heme; Transmembrane.
 FT METAL 96 96 IRON 1 (HEME B562 AXIAL LIGAND).
 FT METAL 110 110 IRON 2 (HEME B566 AXIAL LIGAND).
 FT METAL 197 197 IRON 1 (HEME B562 AXIAL LIGAND).
 FT METAL 211 211 IRON 2 (HEME B566 AXIAL LIGAND).
 SQ SEQUENCE 419 AA; 47237 MW; 202C0C3D25B2A6DB CRC64;
 Query Match 21.2%; Score 7; DB 1; Length 419;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 FLGLGA 14
 DB 351 FLGLGA 357
 RESULT 5
 CG51_HUMAN STANDARD; PRT: 469 AA.
 ID CG51_HUMAN
 AC Q9Y512; Q969Y9; Q96146; Q9NW85; Q9UQM9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Protein CGI-51.
 GN CG151.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANT ILF-345.
 RC MEDLINE=20272150; PubMed=10810093;
 RA Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.;
 RT "Identification of novel human genes evolutionarily conserved in
 Caenorhabditis elegans by comparative proteomics";
 RL Genome Res. 10:703-713(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Kondo H., Sugawara M.,
 RA Nagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Yanamase S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuko Y.,
 RA Ninomiya K., Iwayanagi T.;
 RT "NEDO human cDNA sequencing project";
 RC Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20057165; PubMed=10591208;
 RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
 RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
 RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
 RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
 RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
 RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
 RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
 RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
 RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
 RA Gilbert J.G.R., Goward M.E., Graham D.V., Griffiths M.N.D., Hall C.,
 RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
 RA Hunt S.E., Jones M.C., Kersey J., Kimberley A.M., King A.,
 RA Laird G.K., Langford C.F., Levensha M.A., Lloyd C., Lloyd D.M.,
 RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
 RA McElay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
 RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
 RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
 RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
 RA Soderlund C., Spraggon L., Steward C.A., Sulston J.E., Swann R.M.,
 RA Vaudin K., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
 RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
 RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
 RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
 RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
 RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
 RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Mala J.E., Nguyen T., Pan H.,
 RA Phan Q., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
 RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
 RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,
 RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
 RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,

RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
RA Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,
RA Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R., Willson R.,
RA Emanuel B.S., Shaikh T., Kurahashi H., Saitta S., Budarf M.L.,
RA McDermaid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelman L.,
RA Kim U.J., Shizuya H., Simon M.I., Dumanski C.P., Peyrard M., Kedia D.,
RA Seroussi E., Fransson I., Tapia I., Bruder C.E., O'Brien K.P.,
RA Wilkinson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
RA Tiliahun Y., Wright H.;
RT "The DNA sequence of human chromosome 22.";
RL Nature 402:489-495(1999).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph. Muscle, and Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Schmeier C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore J., Max S.I., Wang C., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz J.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gnaratane P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Houfard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Greenwood J., Schmitt J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska J., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:15899-15903(2002).
CC -!- SIMILARITY: BELONGS TO THE UPF0140 (CGI-51) FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AK001087; BAA91498.1; -
DR EMBL: AF151809; RAD34046.1; -
DR EMBL: AL035398; CAB51401.1; -
DR EMBL: BC007830; AAH07830.1; -
DR EMBL: BC011682; AAH11681.1; -
DR EMBL: BC015200; AAH15200.1; -
DR InterPro: IPR000184; Bac_surfAg_D15.
DR Pfam: PF01103; Bac_surface_Ag; 1.
KW Polymorphism.
FT VARIANT 345 345 V -> I (IN dbSNP:8418).
FT /FTID=VAR_013768.
FT FT
FT CONFLICT 110 110 D -> G (IN REF. 2).
FT CONFLICT 368 372 WAGGL -> IGRRW (IN REF. 1).
FT CONFLICT 371 371 MISSING (IN REF. 4; AAH07830).
SQ SEQUENCE 469 AA: 51962 MW: 4F687D2A12092EF CRC64:
Query Match 21.2%; Score 7; DB 1; Length 469;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 14 ADRFYLG 20
Db 334 ADRFYLG 340
|||||||
RESULT 6
CEAB_ECOLI STANDARD; PRT; 510 AA.
ID CEAB_ECOLI

AC PC5819;
DC 01-NOV-1988 (Rel. 09, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Colicin B.
GN CBA.
OS Escherichia coli.
OG Plasmid ColMB-pPfl66
CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC Enterobacteriaceae; Escherichia.
OX NCB:TaxID=562;
KN [1]
KP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=87250309; PubMed=2439491;
KA Schramm E., Mende J., Braun V., Kamp R.M.;
RT "Nucleotide sequence of the colicin B activity gene cba: consensus
pentapeptide among tonB-dependent colicins and receptors.";
RL J. Bacteriol. 169:3350-3357(1987).
CC -!- FUNCTION: THIS COLICIN IS A CHANNEL-FORMING COLICIN. THIS CLASS OF
TRANSMEMBRANE TOXINS DEPOLARIZE THE CYTOPLASMIC MEMBRANE, LEADING
TO DISSIPATION OF CELLULAR ENERGY.
CC -!- FUNCTION: COLICINS ARE POLYPEPTIDE TOXINS PRODUCED BY AND ACTIVE
AGAINST ESCHERICHIA COLI AND CLOSELY RELATED BACTERIA.
CC -!- MISCELLANEOUS: THIS COLICIN REQUIRES TONB FOR ITS UPTAKE.
CC -!- SIMILARITY: HIGH IN THE N-TERMINAL FIRST 300 AMINO ACID RESIDUES
WITH COLICIN D.
CC -!- SIMILARITY: WITH OTHER CHANNEL FORMING COLICINS, HIGHEST WITH
COLICIN A.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M16816; AAA98063.1; -
DR PIR: A27089; IKEGBB.
DR HSP: P04480; ICOL.
DR InterPro: IPR000293; Channel_colicin.
DR InterPro: IPR003058; Cloacin.
DR InterPro: IPR000531; TonB_boxC.
DR Pfam: PF03515; Cloacin; 1.
DR Pfam: PF01024; Colicin; 1.
DR PRINTS: PR00280; CHANNELCOLICIN.
DR PRINTS: PR01295; CLOACIN.
DR ProDom: PD002657; Channel_colicin; 1.
DR PROSITE: PS00276; CHANNELCOLICIN; 1.
DR PROSITE: PS00430; TONB_DEPENDENT_REC_1; 1.
DR PROSITE: PS01556; TONB_DEPENDENT_REC_2; FALSE_NEG.
KW Antibiotic; Bacteriocin; Plasmid; Transmembrane; TonB box.
FT INIT_MET 0 0
FT SITE 16 23 TONB BOX.
FT TRANSMEM 454 474 POTENTIAL.
FT TRANSMEM 476 496 POTENTIAL.
SQ SEQUENCE 510 AA: 54732 MW: 6E4B972CF19245F1 CRC64:
Query Match 21.2%; Score 7; DB 1; Length 510;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 AVALSIF 8
Db 457 AVALSIF 463
|||||||
RESULT 7
ITRI_SCHPO STANDARD; PRT; 575 AA.
ID ITRI_SCHPO
AC Q10286;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)

```

DT 28-FEB-2003 (Rel. 41, Last annotation update);
DE Myo-inositol transporter 1.
GN ITR1 OR SPAC4F8.15 OR SPAC7D4.G1.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=968 F90;
RX MEDLINE=98228265; PubMed=9560432;
RA Niederberger C., Graub R., Schweingruber A.-M., Fankhauser H.,
RA Rusu M., Politeia M., Edenharter L., Schweingruber M.F.;
RA "Exogenous inositol and genes responsible for inositol transport are
RT required for mating and sporulation in Schizosaccharomyces pombe.";
RL Curr. Genet. 33:255-261(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=1859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Scouras J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin A., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean C.,
RA Mooney P., Moute S., Mungall K., Murphy L., Niblett D., Odeli C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkhardt G., Aert R., Rebben J., Glynnoprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Barzel K., Lander I., Beck A., Leinach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Rayon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong C., Forsburg S.L.,
RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Bartell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RJ Nature 415:871-880(2002).
CC -!- FUNCTION: TRANSPORTER FOR MYO-INOSITOL.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/agreement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X98622; CAAG7211.1; -
DR EMBL: Z94530; CAB11061.1; -
DR EMBL: Z99532; CAB16718.1; -
DR PIR: T43400; T43400.
DR GeneDB: SPombe: SPAC4F8.15; -.
DR InterPro: IPR007114; MIPS.
DR InterPro: IPR005828; Sub_transporter.
DR InterPro: IPR005829; Sug_transporter.
DR InterPro: IPR003663; Sugar_transp.
DR Pfam: PF00083; sugar tr. 1.
DR PRINTS: PR00171; SUGTRANSPORT.
DR TIGRFAMs: TIGR00879; SP; 1.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 2.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; FALSE_NEG.

```

```

KW Transmembrane; Sugar transport; Glycoprotein.
FT DOMAIN 1 86 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 87 107 1 (POTENTIAL).
FT DOMAIN 108 129 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 130 150 2 (POTENTIAL).
FT DOMAIN 151 155 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 157 177 3 (POTENTIAL).
FT DOMAIN 178 186 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 187 207 4 (POTENTIAL).
FT DOMAIN 208 215 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 216 236 5 (POTENTIAL).
FT DOMAIN 237 245 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 247 267 6 (POTENTIAL).
FT DOMAIN 268 349 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 350 376 7 (POTENTIAL).
FT DOMAIN 371 376 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 377 397 8 (POTENTIAL).
FT DOMAIN 398 400 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 401 421 9 (POTENTIAL).
FT DOMAIN 422 441 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 442 462 10 (POTENTIAL).
FT DOMAIN 463 486 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 487 507 11 (POTENTIAL).
FT DOMAIN 508 510 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 511 531 12 (POTENTIAL).
FT DOMAIN 532 575 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 575 575 N-LINKED (GLCNAC...)(POTENTIAL).
SQ SEQUENCE 575 AA; 62757 MW; 3B7C5EFF86C596AE CRC64;

Query Match 21.2%; Score 7; DB 1; Length 575;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 FLGWLGA 14
DE 519 FLGWLGA 525

RESULT 8
LINK2_RAT
ID LINK2_RAT STANDARD PRT: 638 AA.
AC P53670;
CD 01-OCT-1996 (Rel. 34, Created)
CD 01-OCT-1996 (Rel. 34, Last sequence update)
DI 28-FEB-2003 (Rel. 41, Last annotation update)
DE LIM domain kinase 2 (tc 2.7.1.-) (LINK2-2).
GN LINK2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS LINK2A; LINK2B; LINK2C AND LINK2D).
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=95380177; PubMed=7651734;
RA Nunoue K., Ohashi K., Okano I., Mizuno K.;
RT "LINK-1 and LINK-2, two members of a LIM motif-containing protein
RT kinase family.";
RJ Oncogene 11:701-710(1995).
CC -!- FUNCTION: DISPLAYS SPRINE/THREONINE-SPECIFIC PHOSPHORYLATION OF
CC MYELIN BASIC PROTEIN AND HISTONE (MBP) IN VITRO (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=4;
CC Name=LINK2A;
CC ISOID=P53670-1; Sequence=Displayed;
CC Name=LINK2B;
CC ISOID=P53670-2; Sequence=VSP_003128;
CC Name=LINK2C;
CC ISOID=P53670-3; Sequence=VSP_003128, VSP_003129, VSP_003130;
CC Name=LINK2D;
CC ISOID=P53670-4; Sequence=VSP_003128, VSP_003131;
CC -!- TISSUE SPECIFICITY: FOUND IN VARIOUS TISSUES AT MODERATE LEVELS.

```

EXCEPT FOR TESTIS, WHICH SHOWS VERY LOW EXPRESSION.

!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

!- SIMILARITY: Contains 2 LIM zinc-binding domains.

!- SIMILARITY: Contains 1 PDZ/DHR domain.

This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: D31874; BAA06673.1; -
 EMBL: D31875; BAA06674.1; -
 EMBL: D31876; BAA06675.1; -
 EMBL: D31877; BAA06676.1; -
 DR F1R: I78846; I78846.
 DR F1R: I78847; I78847.
 DR F1R: I78848; I78848.
 DR HSP: Q05158; IQLI.
 DR InterPro: IPR001781; LIM.
 DR InterPro: IPR001478; PDZ.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: -PRO02290; Ser_thr_kinase.
 DR Pfam: PFC0412; LIM; 2.
 DR Pfam: PFC0595; PDZ; 1.
 DR Pfam: PFC0069; pkinase; 1.
 DR ProDom: PD000094; LIM; 2.
 DR ProDom: PD000094; LIM; 2.
 DR SMART: SM00132; LIM; 2.
 DR SMART: SM00228; PDZ; 1.
 DR PROSITE: PS00478; LIM_DOMAIN_1; 2.
 DR PROSITE: PS00023; LIM_DOMAIN_2; 2.
 DR PROSITE: PS0106; PDZ; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; FAUSE_NER.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat:
 KW LIM domain; Metal-binding; Zinc; Alternative splicing.
 FT DOMAIN 12 63 LIM 1.
 FT DOMAIN 72 124 LIM 2.
 FT DOMAIN 152 239 PDZ.
 FT DOMAIN 331 608 PROTEIN KINASE.
 FT NP_BIND 337 345 ATP (BY SIMILARITY).
 FT BINDING 360 360 ATP (BY SIMILARITY).
 FT ACT_SITE 451 451 BY SIMILARITY.
 FT VARSPPLIC 1 37 MAALAGERENRCGGNIVPLSCRLYRANEAWHSSC ->
 MGSYLSVPAYFTSRDP (in isoform LIMK2B,
 isoform LIMK2C and isoform LIMK2D).
 FT VARSPPLIC 348 371 /FTIG-VSP_003128.
 VTHKATGKVMVWKEILRCDEETOK -> SWEGSGDSQSHK
 OSDGHEGVNSL (in isoform LIMK2C).
 /FTIG-VSP_003128.
 FT VARSPPLIC 372 638 Missing (in isoform LIMK2C).
 /FTIG-VSP_003130.
 FT VARSPPLIC 165 638 Missing (in isoform LIMK2D).
 /FTIG-VSP_003131.
 FT SEQUENCE 638 AA; 72202 MW; A752D525751BB4F3 CRC64;

Query Match 21.2%; Score 7; DR 1; Length 638;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 0; Gaps 0;

OY 4 ALSFLUG 10
 DB 603 ALSFLUG 609
 |||||

RESULT 9
 AG43_ECOLI STANDARD; PRT: 1039 AA.
 ID AG43_ECOLI
 AC P39180; P75614; P76360; P97241; Q46771;

DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Antigen 43 precursor (AG43) (fluffing protein).
 GN FLU OR B2000.
 GS Escherichia coli.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RI "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97251358; PubMed=9097040;
 RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
 RA Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
 RA Mizubuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
 RA Nashimoto H., Nishic Y., Oshima T., Saito N., Sampei G., Seki Y.,
 RA Sivasubram S., Tagami H., Takeda J., Takemoto K., Wada C.,
 RA Yamamoto Y., Horiuchi T.;
 RI "A 460-kb DNA sequence of the Escherichia coli K-12 genome
 corresponding to the 40.1-50.0 min region on the linkage map.";
 RL DNA Res. 3:379-392(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ML 308-225;
 RA Henderson I.R., Owen P.;
 RL S:Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP PRELIMINARY SEQUENCE OF 53-78.
 RC STRAIN=ML 308-225;
 RX MEDLINE=92929704; PubMed=2661530;
 RA Caifrey P., Owen P.;
 RI "Purification and N-terminal sequence of the alpha subunit of antigen
 43, a unique protein complex associated with the outer membrane of
 Escherichia coli.";
 RL J. Bacteriol. 171:3634-3640(1989).
 RN [5]
 RP SEQUENCE OF 53-63.
 RC STRAIN=K12 / EMG2;
 RX MEDLINE=97443975; PubMed=9298646;
 RA Link A.J., Robison K., Church G.M.;
 RI "Comparing the predicted and observed properties of proteins encoded
 in the genome of Escherichia coli K-12.";
 RL Electrophoresis 18:1259-1313(1997).
 RN [6]
 RP GENE NAME.
 RC MEDLINE=97257509; PubMed=9103983;
 RA Henderson I.R., Meehan M., Owen P.;
 RI "Antigen 43, a phase-variable bipartite outer membrane protein,
 determines colony morphology and autoaggregation in Escherichia coli
 K-12.";
 RL FEMS Microbiol. Lett. 149:115-120(1997).
 CC !- FUNCTION: CONTROLS COLONY FORM VARIATION AND AUTOAGGREGATION. MAY
 CC FUNCTION AS AN ADHESIN.
 CC !- SUBUNIT: HETERODIMER OF THE PERIPHERAL MEMBRANE PROTEIN (ALPHA
 CC CHAIN) ANCHORED TO THE INTEGRAL OUTER MEMBRANE PROTEIN (BETA
 CC CHAIN).
 CC !- SUBCELLULAR LOCATION: Outer membrane-associated.
 CC !- SIMILARITY: TO ADHESIN AIDA-I AND TO BORDETTELLA PERTACTIN.
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AE000291; AAC75061.1; ALT_INIT.
DR EMBL: D90838; BAA15825.1; ALT_INIT.
DR EMBL: D90839; BAA15832.1; ALT_INIT.
DR EMBL: U24429; AAB47869.1; -.
DR HSSP: PG7505; LSND.
DR EcoGene: EG12686; flu.
DR InterPro: IPR000635; Autotransporter.
DR InterPro: IPR005546; Autotransporter.
DR InterPro: IPR004899; Pertactin.
DR Pfam: PF03797; Autotransporter; 1.
DR Pfam: PF03212; Pertactin; 1.
DR TIGRFAMs: TIGR01414; autotrans_bar1; 1.
KW Outer membrane; Signal; Complete proteome.
FT SIGNAL 1 52
FT CHAIN 53 551 ANTIGEN 43 ALPHA CHAIN.
FT CHAIN 552 1039 ANTIGEN 43 BETA CHAIN.
FT VARIANT 2 2 SL -> N (IN STRAIN ML 308-225).
FT VARIANT 41 42 SL -> FF (IN STRAIN ML 308-225).
FT VARIANT 46 46 T -> K (IN STRAIN ML 308-225).
FT VARIANT 157 157 W -> L (IN STRAIN ML 308-225).
FT VARIANT 188 188 V -> F (IN STRAIN ML 308-225).
FT VARIANT 303 305 AIN -> STI (IN STRAIN ML 308-225).
FT VARIANT 320 320 A -> T (IN STRAIN ML 308-225).
FT VARIANT 372 372 N -> Q (IN STRAIN ML 308-225).
FT VARIANT 493 493 E -> V (IN STRAIN ML 308-225).
FT VARIANT 497 497 S -> N (IN STRAIN ML 308-225).
FT VARIANT 585 585 H -> Y (IN STRAIN ML 308-225).
FT VARIANT 709 709 E -> K (IN STRAIN ML 308-225).
FT VARIANT 721 721 M -> T (IN STRAIN ML 308-225).
FT VARIANT 751 753 GHL -> SHF (IN STRAIN ML 308-225).
FT VARIANT 803 603 S -> P (IN STRAIN ML 308-225).
FT VARIANT 815 615 A -> V (IN STRAIN ML 308-225).
FT VARIANT 824 824 C -> S (IN STRAIN ML 308-225).
FT VARIANT 829 835 LNLVHTS -> MNLVYNA (IN STRAIN ML 308-225).
FT VARIANT 845 847 QGT -> LGA (IN STRAIN ML 308-225).
FT VARIANT 855 855 S -> T (IN STRAIN ML 308-225).
FT VARIANT 888 888 Q -> L (IN STRAIN ML 308-225).
FT VARIANT 1025 1025 S -> I (IN STRAIN ML 308-225).
FT CONFLICT 61 63 EIV -> IIT (IN REP. 5).
SQ SEQUENCE 1039 AA; 10584; MW: 51706647C8DEFEBC CRC64;

Query Match 21.2%; Score 7; DB 1; Length 1039;
Best Local Similarity 100.0%; Pred No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAVALSJ 7
Db 36 VAVALSJ 42

RESULT 10
PSAG_PEA STANDARD; PRT; 39 AA.
AC P20120;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Photosystem I reaction center subunit V (PSI-G) (Photosystem I 9 kDa protein) (Fragment).
DE PSAG.
GN Pisum sativum (Garden pea).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE.

RX MEDLINE=89137587; PubMed=3277857;
RA Dunn P.J., Packman L.C., Pappin D., Gray J.C.;
RT "N-terminal amino acid sequence analysis of the subunits of pea
photosystem I";
RL FEBS Lett. 228:157-161(1988).
CC -!- FUNCTION: NOT YET KNOWN.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast
thylakoid membrane (Probable).
CC -!- SIMILARITY: Belongs to the psag / psak family.
DE PIR: S00318; S00318.
DE InterPro: IPR000549; PSI_Psag/K.
DE Pfam: PF01241; PSI_Psag; 1.
DE PROSITE: PS01026; PHOTOSYSTEM_I_Psag; PARTIAL.
KW Photosynthesis; Photosystem I; Transmembrane; Chloroplast; Thylakoid.
FT NON_TER 39 39
SQ SEQUENCE 39 AA; 4397 MW; 00EE8A9E219D91A3 CRC64;

Query Match 18.2%; Score 6; DB 1; Length 39;
Best Local Similarity 100.0%; Pred No. 19;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LSLFLG 10
Db 14 LSLFLG 19

RESULT 11
TY13_HUMAN STANDARD; PRT; 58 AA.
ID TY13_HUMAN
AC Q9BZ97;
DI 28-FEB-2003 (Rel. 41, Created)
DI 28-FEB-2003 (Rel. 41, Last sequence update)
DI 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transcript Y13 protein.
GN TY13 OR TY13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue-Testis;
KA Kuroda-Kawaguchi T., Skaletsky H., Minx P.J., Brown L.G., Rozen S.,
KA Wilson R.K., Waterston R.H., Page D.C.;
RC "The DNA sequence of the human Y chromosome";
RC Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AF332242; AAK13492.1; -.
DS Genew; HGNC:18494; TTY13.
KW Transmembrane.
FT TRANSMEM 17 37 POTENTIAL.
SQ SEQUENCE 58 AA; 6256 MW; F714A679F062DFE; CRC64;

Query Match 18.2%; Score 6; DB 1; Length 58;
Best Local Similarity 100.0%; Pred No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LSLFLG 10
Db 24 LSLFLG 29

RESULT 12
YG12_BACHD STANDARD; PRT; 111 AA.
ID YG12_BACHD

AC Q9KCF9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein BH1612.
 GN BH1612.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A. 9153;
 RX STRAIN=C-125 / JCM 9153;
 RC MEDLINE=20512582; PubMed=11088132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirata C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis."
 RL Nucleic Acids Res. 28:4317-4331(2000).
 CC -!- SIMILARITY: BELONGS TO THE UPF0131 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AP001512; BAB05331.1;
 DR PIR: D83851; D83851.1;
 DR InterPro: IPR005347; JPF0131.
 DR Pfam: PF03674; UPF0131.1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 111 AA: 12782 MW: 24730C79CC2096FE CRC64;

 Query Match 18.28; Score 6; DB 1; Length 111;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 19 LGYPAL 24
 Db 22 LGYPAL 27

 RESULT 13
 ID PSCA_HUMAN STANDARD; PRT; 123 AA.
 AC O43653;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Prostate stem cell antigen precursor.
 GN PSCA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Prostatic carcinoma;
 RC MEDLINE=98132661; PubMed=9465086;
 RA Reiter R.F., Gu Z., Watabe T., Thomas G., Szigeti K., Davis E.,
 RA Wahl M., Nistiani S., Yamashiro J., le Beau M.M., Loda M., Witte O.N.;
 RT "Prostate stem cell antigen: a cell surface marker overexpressed in
 RT prostate cancer."
 RL Proc. Natl. Acad. Sci. U.S.A. 95:1735-1740(1998).
 CC [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Urothelial;
 RA Bahrenberg G., Joost H.G.;
 RT "Serial analysis of the gene expression of a highly differentiated
 RT urothelial tumor."

KL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=20180504; PubMed=10713670;
 RA Gu Z., Thomas G., Yamashiro J., Shintaku I.P., Dorey F., Raitano A.,
 RA Witte O.N., Said J.W., Loda M., Reiter R.F.;
 RT "Prostate stem cell antigen (PSCA) expression increases with high
 RT gleason score, advanced stage and bone metastasis in prostate
 RT cancer."
 RL Oncogene 19:1288-1296(2000).
 CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PROSTATE (BASAL, SECRETORY
 CC AND NEUROENDOCRINE EPITHELIUM CELLS). ALSO FOUND IN BLADDER
 CC (TRANSITIONAL EPITHELIUM), PLACENTA (TROPHOBLASTS), STOMACH
 CC (NEUROENDOCRINE CELLS), COLON (NEUROENDOCRINE CELLS) AND KIDNEY
 CC (COLLECTING DUCTS); OVEREXPRESSED IN PROSTATE CANCERS AND
 CC EXPRESSION IS CORRELATED WITH TUMOR STAGE, GRADE AND ANDROGEN-
 CC INDEPENDENCE. HIGHLY EXPRESSED IN PROSTATE CANCER BONE METASTASES.
 CC -!- SIMILARITY: Contains 1 UPAR/Ly6 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF043498; AAC39657.1;
 DR EMBL: AJ297436; CAB97347.1;
 DR GenBank: HGNC:9500; PSCA.
 DR MIM: 602470;
 DR GO: GO:0005886; C:plasma membrane; TAS.
 DR GO: GO:0008222; F:tumor antigen; TAS.
 DR InterPro: IPR003632; Ly-6_CD59.
 DR InterPro: IPR001526; Ly6_UPAR.
 DR Pfam: PF00021; UPAR_Ly6; 1.
 DR ProDom: PD003128; Ly-6_CD59; 1.
 DR SMART: SM00134; LU; 1.
 DR PROSITE: PS00983; Ly6_UPAR; FALSE NEG.
 KW Signal; Antigen; Glycoprotein; Membrane; GPI-anchor.
 FT SIGNAL 1 20 POTENTIAL
 FT CHAIN 21 95 PROSTATE STEM CELL ANTI-GEN.
 FT PROPEP 96 123 REMOVED IN MATURE FORM (POTENTIAL).
 FT DOMAIN 21 95 UPAR/LY6.
 FT DISULFID 23 48 BY SIMILARITY.
 FT DISULFID 26 35 BY SIMILARITY.
 FT DISULFID 41 66 BY SIMILARITY.
 FT DISULFID 70 86 BY SIMILARITY.
 FT DISULFID 87 92 BY SIMILARITY.
 FT CARBOHYD 40 40 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 83 83 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 93 93 N-LINKED (GLCNAC...) (POTENTIAL).
 FT LIPID 95 95 GPI-ANCHOR (POTENTIAL).
 SQ SEQUENCE 123 AA: 12912 MW: 3FC1271742D657FA CRC64;

 Query Match 18.28; Score 6; DB 1; Length 123;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 22 PALGLL 27
 Db 111 PALGLL 116

 RESULT 14
 ID GTH2_MORSA STANDARD; PRT; 139 AA.
 AC Q91121;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Gonadotropin beta-II chain precursor (GTH-II-beta).

OS Morone saxatilis (Striped bass).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
 OC Moronidae; Morone.
 OX NCBI_TaxID=34816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pituitary;
 RX MEDLINE=96020549; PubMed=8546811;
 RA Hassin S., Elizur A., Zohar Y.;
 RI "Molecular cloning and sequence analysis of striped bass (Morone
 saxatilis) gonadotropin-I and -II subunits.";
 RL J. Mol. Endocrinol. 15:23-35(1995).
 CC -!- FUNCTION: INVOLVED IN GAMETOGENESIS AND STEROIDOGENESIS.
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta chain.
 CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L35096; AAC38019.1; ALT_INIT.
 CC HSSP: P01233; 1XUL.
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR002400; GF_CysKnot.
 DR InterPro: IPR001545; Gly_hormoneB.
 DR Pfam: PF00007; Cys_knot; 1.
 DR PRINTS: PR03438; GFCYSKNOT.
 DR SMART: SM00068; GHF; 1.
 DR PROSITE: PS00261; GLYCO_HORMONE_BETA_1; 1.
 DR PROSITE: PS00689; GLYCO_HORMONE_BETA_2; 1.
 KW Hormone; Glycoprotein; Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 139 GONADOTROPIN BETA-1 CHAIN.
 FT DISULFID 30 78 BY SIMILARITY.
 FT DISULFID 44 93 BY SIMILARITY.
 FT DISULFID 47 131 BY SIMILARITY.
 FT DISULFID 55 109 BY SIMILARITY.
 FT DISULFID 59 111 BY SIMILARITY.
 FT DISULFID 114 121 BY SIMILARITY.
 FT CARBOHYD 34 34 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 139 AA; 15555 MW; 1AB8B742FB81AF76 CRC64;
 Query Match: 18.2%; Score 6; DB 1; Length 139;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 LSLFLG 10
 Db 11111
 6 LSLFLS 11
 RESULT 15
 PSAG_ARATH
 ID PSAG_ARATH STANDARD; PRT; 160 AA.
 AC Q95N7; Q42310;
 DT 28-FEB-2003 (Rel. 41; Created)
 DT 28-FEB-2003 (Rel. 41; Last sequence update)
 DT 28-FEB-2003 (Rel. 41; Last annotation update)
 DE Photosystem I reaction center subunit v, chloroplast precursor (PS-
 G).
 DE PSAG OR ATIG55670 OR F20N2.3 OR F20N2.33 OR F20N2.21.
 GN Arabidopsis thaliana (Mouse-ear cress).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 Query Match: 18.2%; Score 6; DB 1; Length 160;

[1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=cv. Columbia;
 RC "Legen J., Misera S., Herrmann R.G., Altschmied E.;
 *Sequences and map position of 31 Arabidopsis thaliana cDNAs encoding
 organellar polypeptides.";
 RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=22016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
 Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 Dunn P., Egtu P., Feildlyum T.V., Feng J.-B., Fong B., Fujii C.Y.,
 Giller J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,
 Miittscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana.";
 RL Nature 408:816-820(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 "Riken Arabidopsis full length cDNA clones (RAFLs) sequenced by the
 SSP consortium (Salk/Stanford/PGEC).";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 27-127 FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Bardet C., Dabos P., Tremoussaygue D., Lescure B.;
 Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: NOT YET KNOWN.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast
 CC thylakoid membrane (Probable).
 CC -!- SIMILARITY: Belongs to the psag / psak family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AJ245630; CAB52748.1; -
 CC EMBL: AC002328; AAC83069.1; -
 CC EMBL: AF326870; AAC41452.1; -
 CC EMBL: AF324710; AAC40061.1; -
 CC EMBL: AF339692; AAK00374.1; -
 CC EMBL: Z37244; CAA85530.1; -
 CC In-crPro: IPRC00549; PSI_Psag/K.
 CC Pfam: PF01241; PSI_PsAK; 1.
 DR PROSITE: PS01026; PHOTOSYSTEM_I_PSAGK; 1.
 KW Photosynthesis; Photosystem I; transmembrane; Chloroplast; Thylakoid;
 KW Transit peptide.
 FT TRANSIT 1 60 CHLOROPLAST (BY SIMILARITY).
 FT CHAIN 61 160 PHOTOSYSTEM I REACTION CENTER SUBUNIT V.
 FT TRANSMEM 66 86 POTENTIAL.
 FT TRANSMEM 129 149 POTENTIAL.
 SQ SEQUENCE 160 AA; 17085 MW; 6EF23E57C7C50760 CRC64;
 Query Match: 18.2%; Score 6; DB 1; Length 160;

Best Local Similarity 100.0%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LSLFLG 10
jlllll
Db 74 LSLFLG 79

Search completed: September 26, 2003, 17:37:20
Job time : 4.93192 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:29:45 ; Search time 17.1319 seconds
(without alignments)
497.069 Million cell updates/sec

Title: US-09-852-100A-2_COPY_185_217
Perfect score: 33
Sequence: 1 VAVALSFLGWLGWGADRFYLGYPALGLIKFCVVG 33

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first: 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	100.0	207	4 Q9BX74	Q9BX74 Homo sapiens
2	33	100.0	208	11 Q99MB3	Q99MB3 Mus musculus
3	8	24.2	80	4 Q9H046	Q9H046 Homo sapiens
4	8	24.2	178	5 Q9W2H1	Q9W2H1 Drosophila
5	8	24.2	195	5 Q95Q25	Q95Q25 Caenorhabditis
6	8	24.2	221	4 Q9H651	Q9H651 Homo sapiens
7	8	24.2	230	11 Q9D156	Q9D156 Mus musculus
8	8	24.2	247	4 Q9BRN9	Q9BRN9 Homo sapiens
9	8	24.2	261	11 Q8BJ83	Q8BJ83 Mus musculus
10	8	24.2	284	5 Q9H4H5	Q9H4H5 Drosophila
11	7	21.2	38	11 Q8CAM1	Q8CAM1 Mus musculus
12	7	21.2	150	16 Q8EV17	Q8EV17 Mycoplasma
13	7	21.2	159	10 Q9LXV9	Q9LXV9 Arabidopsis
14	7	21.2	180	16 Q8FPV7	Q8FPV7 Corynebacterium
15	7	21.2	194	16 Q92QW4	Q92QW4 Rhizobium
16	7	21.2	219	16 Q9RX10	Q9RX10 Deinococcus

17	7	21.2	221	12 Q91TN3	Q91TN3 Tupia herp
18	7	21.2	251	11 Q8BS94	Q8BS94 Mus musculus
19	7	21.2	259	17 Q8PVR4	Q8PVR4 Methanosaeta
20	7	21.2	269	16 Q8RRR3	Q8RRR3 Thermoplasma
21	7	21.2	273	16 Q83115	Q83115 Treponema
22	7	21.2	314	2 Q8RLA5	Q8RLA5 Escherichia
23	7	21.2	317	17 Q57894	Q57894 Pyrococcus
24	7	21.2	329	12 Q8U2J1	Q8U2J1 Cercopithecus
25	7	21.2	349	16 Q9PLE9	Q9PLE9 Chlamydia
26	7	21.2	358	16 Q8CUNG	Q8CUNG Oceanobacillus
27	7	21.2	401	2 Q9S458	Q9S458 Salmonella
28	7	21.2	407	16 Q8ED47	Q8ED47 Shewanella
29	7	21.2	415	16 Q915C8	Q915C8 Pseudomonas
30	7	21.2	444	2 Q9RQ60	Q9RQ60 Rhizobium
31	7	21.2	448	2 Q24842	Q24842 Acinetobacter
32	7	21.2	451	11 Q9QU14	Q9QU14 Mus musculus
33	7	21.2	466	2 Q87185	Q87185 Streptococcus
34	7	21.2	466	2 Q9ALW7	Q9ALW7 Streptococcus
35	7	21.2	466	2 Q9AFH1	Q9AFH1 Streptococcus
36	7	21.2	466	2 Q93113	Q93113 Streptococcus
37	7	21.2	466	2 Q9AQ17	Q9AQ17 Streptococcus
38	7	21.2	466	16 Q8E500	Q8E500 Streptococcus
39	7	21.2	466	16 Q8D2E2	Q8D2E2 Streptococcus
40	7	21.2	469	11 Q8BGH2	Q8BGH2 Mus musculus
41	7	21.2	492	16 Q9RUA0	Q9RUA0 Deinococcus
42	7	21.2	510	17 Q59101	Q59101 Pyrococcus
43	7	21.2	547	16 Q9R2J4	Q9R2J4 Deinococcus
44	7	21.2	572	10 Q8VYV8	Q8VYV8 Arabidopsis
45	7	21.2	606	13 Q8QHMO	Q8QHMO Xenopus

ALIGNMENTS

RESULT 1

Q9BX74 ID Q9BX74 PRELIMINARY: PRT: 207 AA.
AC Q9BX74;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DI 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Beta-amyloid binding protein.
GN BAP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA Kajkowski E.M., Lo C.F., Ning X., Walker S., Sofia H.J., Wang W.,
Edris W., Chanda P., Wagner E., Vile S., Ryan K., McHenry-Rinde H.,
Smith S.C., Wood A., Rhodes K.J., Kennedy J.D., Bard J.,
Jacobsen J.S., Ozenberger B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein.
Containing a G Protein Activation Module.";
RE J. Biol. Chem. 275:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA Kajkowski E.M., Lo C.F., Ning X., Walker S., Sofia H.J., Wang W.,
Edris W., Chanda P., Wagner E., Vile S., Ryan K., McHenry-Rinde H.,
Smith S.C., Wood A., Rhodes K.J., Kennedy J.D., Bard J.,
Jacobsen J.S., Ozenberger B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein.
Containing a G Protein Activation Module.";
RE J. Biol. Chem. 275:18748-18756(2001).

Query Match 100.0%; Score 33; DB 4; Length 207;
Best Local Similarity 100.0%; Pred. No. 2.1e-24;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Q9 1 VAVALSFLGWLGWGADRFYLGYPALGLIKFCVVG 33

Db 123 VAVALSLFLWLGADRFYLGYPALGLLKFTCVG 155
|||||

RESULT 2

Q99MB3 PRELIMINARY; PRT; 208 AA.
AC Q99MB3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Beta-amyloid binding protein.
GN BBP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=21276355; PubMed=11278849;
RA Kajkowski E.M., Lo C.F., Ning X., Walker S., Sofia H.J., Wang W.,
RA Edris W., Chanda P., Wagner E., Vile S., Ryan K., McHenry-Rinde E.,
RA Smith S.C., Wood A., Rhodes K.J., Kennedy J.D., Bard J.,
RA Jacobsen J.S., Ozenberger B.A.,
RT "beta-Amyloid Peptide-Induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RL J. Biol. Chem. 276:18748-18756(2001).
DR EMBL: AF353993; AK35067.1;
DR MGD: MG1:2137022; Bbp.
SQ SEQUENCE 208 AA; 22271 MW; 92A7932163F4F04C CRC64;
Query Match 100.0%; Score 33; DB 11; Length 268;
Best Local Similarity 100.0%; Pred. No. 2.2e-24;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VAVALSLFLWLGADRFYLGYPALGLLKFTCVG 33
|||||

Db 124 VAVALSLFLWLGADRFYLGYPALGLLKFTCVG 156
|||||

RESULT 3

Q9H046 PRELIMINARY; PRT; 80 AA.
AC Q9H046;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein: (fragment).
GN DKFZP667C1011.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph node;
RA Koehler K., Beyer A., Mewes H.W., Weil R., Wiemann S.,
RL Submitted (DBC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL512689; CAC21647.1;
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 80 AA; 8699 MW; 8B86BE788235C58D CRC64;
Query Match 24.2%; Score 8; DB 4; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 GADRFYLG 20
|||||

Db 33 GADRFYLG 40
|||||

RESULT 4

Q9W2H1 PRELIMINARY; PRT; 178 AA.
AC Q9W2H1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CG10795 protein (LD27358P).
GN CG10795.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bereman B.P., Bhardwaj D., Bolshakov S.,
RA Borokova J., Botchan M.R., Bouck J., Brockstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Fertaz C., Ferreira S., Fleischmann W.,
RA Fessler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.D., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Keichum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
RA Patzoldt M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun F.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brockstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuncio J., Paclab J., Parag V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.,
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF003453; AAF46720.1;
DR EMBL: AY061343; AAL28891.1;
DR Flybase: FBgn0034626; CG10795.
SQ SEQUENCE 178 AA; 19896 MW; 17C41166607ACC03 CRC64;
Query Match 24.2%; Score 8; DB 5; Length 178;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Caps 0;

```

QY      15 DREYLYGP 22
      1111111
DB      108 DREYLYGP 115

RESULT 5
Q95Q25 PRELIMINARY: PRT: 195 AA.
AC Q95Q25;
DT 01-DEC-2001 (TrEMBLrel. 19, Created);
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update);
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update);
DE Hypothetical 21.2 kDa protein.
GN C41D11.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peziderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=95069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gattug S., Magdi L.;
RT "The sequence of C. elegans cosmid C41D11.";
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF003740; AAL08031.1;
DR WormPeP: C41D11.9; CE29489.
KW Hypothetical protein.
SQ SEQUENCE 195 AA: 21203 MW: 35945E407F134DAE CRC64;
      24.2% Score 8; DB 5; Length 195;
      Best Local Similarity 100.0%; Pred. No. 4.8;
      Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      13 GADRYFLG 20
      1111111
DB      148 GADRYFLG 155

RESULT 6
Q9H651 PRELIMINARY: PRT: 221 AA.
AC Q9H651;
DT 01-MAR-2001 (TrEMBLrel. 16, Created);
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update);
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update);
DE Hypothetical protein FLJ22604 (BBP-like protein 2).
GN BLP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isozaki T., Sugano S.;
RT "NEO human cDNA sequencing project.";

```

```

RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kajkowski E.M., Lo C.F., Ning X., Walker S., Sofia H.J., Wang W.,
RA Edris W., Charada P., Wagner E., Ville S., Ryan K., McHenry-Rinde B.,
RA Smith S.C., Wood A., Rhodes K.J., Kennedy J.D., Bard J.,
RA Jacobsen J.S., Gzenberger B.A.;
RC "Beta-amyloid peptide-induced apoptosis regulated by a novel protein
RC containing a G protein activation module.";
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK026257; BAB15415.1;
DR EMBL: AF353992; AAK35066.1;
DR EMBL: BC008873; AAH08873.1;
KW Hypothetical protein.
SQ SEQUENCE 221 AA: 24410 MW: 92151D6EF6363D74 CRC64;
      24.2% Score 8; DB 4; Length 221;
      Best Local Similarity 100.0%; Pred. No. 5.3;
      Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      13 GADRYFLG 20
      1111111
DB      174 GADRYFLG 181

RESULT 7
Q9D156 PRELIMINARY: PRT: 230 AA.
AC Q9D156;
DT 01-JUN-2001 (TrEMBLrel. 17, Created);
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update);
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update);
DE 1110025109rik protein (RIKEN cDNA 1110025109 gene).
GN 1110025109RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RC MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gotojori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Frieschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schram L.M., Staubil F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Morgone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilmink L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kotsuki S.,
RA Hayashizaki Y.;
RC "Functional annotation of a full-length mouse cDNA collection.";
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK003917; BAB23075.1;

```

DR EMBL BC024620: AAH24620.1: -
 DR MGD: MGI:1915884: 1110025109Rix.
 SQ SEQUENCE 230 AA: 25639 MW: 396D650D8BE99A5 CRC64;
 Query Match 24.2% Score 8: DB 11: Length 230;
 Best Local Similarity 100.0%; Pred. No. 5.5;
 Matches 8: Conservative 0: Mismatches 0: Indels 0: Gaps 0;

QY 13 GADRFYLG 20
 |||||
 DB 183 GADRFYLG 190

RESULT 8
 Q9BRN9 PRELIMINARY; PRT: 247 AA.
 AC Q9BRN9: 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-DISC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Similar to hypothetical protein F1J22604.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Colon;
 RA Strausberg R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC006150; AAH06150.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 247 AA: 27161 MW: CE1D0D5C53DF73C CRC64;
 Query Match 24.2% Score 8: DB 4: Length 247;
 Best Local Similarity 100.0%; Pred. No. 5.9;
 Matches 8: Conservative 0: Mismatches 0: Indels 0: Gaps 0;

QY 13 GADRFYLG 20
 |||||
 DB 200 GADRFYLG 207

RESULT 9
 Q8BJ83 PRELIMINARY; PRT: 261 AA.
 AC Q8BJ83:
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DE Similar to BBP-like protein 2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Forelimb;
 RX MEDLINE=22354583; PubMed=12466851;
 RA The FANTOM Consortium.
 RA The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs.
 RL Nature 420:563-573(2002).
 DR EMBL: AK077858; BAC37037.1; -
 SQ SEQUENCE 261 AA: 28880 MW: 70346780D3CF5CDB CRC64;
 Query Match 24.2% Score 8: DB 11: Length 261;
 Best Local Similarity 100.0%; Pred. No. 6.1;
 Matches 8: Conservative 0: Mismatches 0: Indels 0: Gaps 0;

QY 13 GADRFYLG 20
 |||||

DB 214 GADRFYLG 221

RESULT 10
 Q9U4H5 PRELIMINARY; PRT: 284 AA.
 AC Q9U4H5: 09M361:
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE BCDNA:GH02974 (ALMONDEX) (AMX protein).
 GN AMX OR BCDNA:GH02974 OR CG12127.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha.
 NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
 RA Agbayani A., Arcaina T.T., Baxter E., Blazej R.G., Butenhoif C.,
 RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
 RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
 RA Houston K.A., Hummasti S.R., Kim E., Li P., Moshrefi M., Pacleb J.M.,
 RA Park S., Sequeira A., Sethi H., Snir E., Svirska R.R., Weinburg T.,
 RA Celisner S.E.;
 RT "Full Length Drosophila melanogaster cDNA sequence."
 RC Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RA Micheliou M.A.E., Remillieux N.C., Randsholt N.B.;
 RT "Characterization of almondex."
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.K., Bouck J., Brokstein P., Brottier P.,
 RA Surtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Durkin B.C., Dun P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Calali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Minkov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou S., Zhu X., Smith H.O.,

RA Gibos R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 DR EMBL: AF181623; AAD53409.1; -.
 DR EMBL: AF211797; AAF36924.2; -.
 DR EMBL: AE003446; AAF46474.2; -.
 DR FlyBase: FBgn0000077; amx.
 DR InterPro: IPR001304; Lectin_C.
 DR PROSITE: PS00615; C_TYPE_LECTIN_1; 1.
 SQ SEQUENCE 284 AA; 31364 MW; 8FB8FF573AC851 CRC64;

Query Match 24.2%; Score 8; DB 5; Length 284;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 GADRFYLG 20
 DB 237 GADRFYLG 244

RESULT 11

Q8CAM1 ID Q8CAM1 PRELIMINARY; PRT; 38 AA.
 AC Q8CAM1;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE LIM motif-containing protein kinase 2 (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Hypothalamus;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium.
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 DR EMBL: AK038530; BAC30030.1; -.
 FT NON_TER 1
 SQ SEQUENCE 38 AA; 4188 MW; 2DB363A494415D42 CRC64;

Query Match 21.2%; Score 7; DB 11; Length 38;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ALSFLG 10
 DB 3 ALSFLG 9

RESULT 12

Q8EVI7 ID Q8EVI7 PRELIMINARY; PRT; 150 AA.
 AC Q8EVI7;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN MYPE5770.
 OS Mycoplasma penetrans.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=28227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HF-2;
 RX MEDLINE=22354719; PubMed=12466555;
 RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,
 RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;
 RT "The complete genomic sequence of *Mycoplasma penetrans*, an

RT Intracellular bacterial pathogen in humans."
 RL Nucleic Acids Res. 30:5293-5300(2002).
 DR EMBL: AP004172; BAC44367.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 150 AA; 17289 MW; 0FF392EB9F32F1F1 CRC64;

Query Match 21.2%; Score 7; DB 16; Length 150;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 GADRFYLG 19
 DB 70 GADRFYLG 76

Q5LXV9 ID Q5LXV9 PRELIMINARY; PRT; 159 AA.
 AC Q5LXV9;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN T15a3.160 OR AT3G44020/T15B3.160.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Jordan N., Bangert S., Wiedemann R., Voss W., Unseld M., Mewes H.W.,
 RA Rudd S., Lemcke K., Mayer K.F.X., Quettner F., Salanoubat M.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
 RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
 RA Hayashizaki Y., Shinozaki K.;
 RT "Arabidopsis thaliana full-length cDNA."
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL163975; CAB88130.1; -.
 DR EMBL: AK118643; BAC43239.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 159 AA; 17059 MW; A3F0A7B03B5AR78A CRC64;

Query Match 21.2%; Score 7; DB 10; Length 159;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ALSFLG 10
 DB 94 ALSFLG 100

RESULT 14

Q8FPV7 ID Q8FPV7 PRELIMINARY; PRT; 180 AA.
 AC Q8FPV7;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Conserved hypothetical protein.
 GN CE1361.
 OS Corynebacterium efficiens.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=152794;

RN SEQUENCE FROM N.A.
 RP STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
 RA Kawarabayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
 RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
 RA Usuda Y., Sugimoto S.,
 RA "The entire genomic sequence of *Corynebacterium efficiens* YS-314."
 RT Submitted (MAY-2002) to the EMBL/GenBank/DBS databases.
 RL EMBL: AP005218; BAC18191.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 180 AA; 19455 MW; EA6BF8CA54FCD761 CRC64;

Query Match 21.2%; Score 7; DB 16; Length 180;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 PALGLLK 28
 |||||
 Db 133 PALGLLK 139

RESULT 15

O92QW4
 ID O92QW4 PRELIMINARY: PRT: 194 AA.
 AC O92QW4:
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein acvB.
 GN ACV8 OR R01182 OR SMC00613.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID:382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396507; PubMed=11481430;
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Barut J.,
 RA Boistard P., Becker A., Boutry M., Cadieu E., Drearo S., Gloux S.,
 RA Godrie T., Coffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
 RA Pohl T., Portetelle D., Pehier A., Purnelle B., Ramsperger U.,
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.,
 RT "Analysis of the chromosome sequence of the legume symbiont
 RT Sinorhizobium meliloti strain 1021."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 DR EMBL: AL591786; CAC45761.1;
 DR InterPro; IPR000734; Lipase.
 DR InterPro; IPR000379; Ser_estrs_site.
 DR PROSITE; PS00120; LIPASE_SER; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 194 AA; 21357 MW; AABE97B2813F5E42 CRC64;

Query Match 21.2%; Score 7; DB 16; Length 194;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 LGWLGAD 15
 |||||
 Db 113 LGWLGAD 119

Search completed: September 26, 2003, 17:41:38
 Job time : 20.1319 secs

XX
PI Bard JA, Jacobsen JS, Kajkowski EM, Ozenberger BA;
PI Walker SG;

```

XX WPI: 1999-080736/07.
DR N-PSDB: AAX05735.
XX
PT Polynucleotide encoding beta-amyloid peptide binding protein - used
PT to identify inhibitors of beta-amyloid peptide for treating
PT Alzheimer's disease
XX
XX Claim 7: Pages 43-44; 59pp; English.
XX
CC The present sequence represents a beta-amyloid peptide binding protein
CC (BBP). The polynucleotide comprising the entire BBP nucleotide sequence
CC of clone BBP1-fl is deposited under the accession number ATCC 98617. The
CC polynucleotide comprising a fragment of BBP (nucleotides 202-807 of the
CC full length BBP) of clone pBK196 is deposited as ATCC 98399. Host cells
CC transformed with a vector comprising the BBP nucleic acid are used in a
CC method for diagnosing a disease characterised by aberrant expression of
CC human beta-amyloid protein (BAP). The protein can also be used in a
CC method for screening for compounds which regulate expression of a BAP
CC binding protein. The proteins, antibodies and identified compounds can be
CC used in the treatment or prevention of Alzheimer's disease.
XX
XX Sequence 269 AA:
XX
XX Query Match 100.0%; Score 202; DB 20; Length 269;
XX Best Local Similarity 100.0%; Pred. No. 7e-193;
XX Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 PSGPSAPEAVTARLVGVLFVSVITGPGAVATSGAGEESLKCEDLKVGQYICKDKPKIND 60
Db 68 PSGPSAPEAVTARLVGVLFVSVITGPGAVATSGAGEESLKCEDLKVGQYICKDKPKIND 127
XX
QY 61 ATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFKPISCRNVNGYSYKVAV 120
Db 128 ATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFKPISCRNVNGYSYKVAV 187
XX
QY 121 ALSFLGWLGAADRFYLGYPALGLKFCITGFCIGSLIDFILISMQIVGPSDGSSTYIDY 180
Db 188 ALSFLGWLGAADRFYLGYPALGLKFCITGFCIGSLIDFILISMQIVGPSDGSSTYIDY 247
XX
QY 181 YGTRLTRLSITNETFRKTQLYP 202
Db 248 YGTRLTRLSITNETFRKTQLYP 269
XX
XX RESULT 2
XX AAY70759
XX ID AAY70759 standard; Protein: 269 AA.
XX AC AAY70759;
XX DT 24-JUL-2000 (first entry)
XX DE Human beta-amyloid peptide (BAP) binding protein, BBP1.
XX
KW Beta-amyloid peptide binding protein; BBP; BAP; tumour; suppressor;
KW G-protein coupled receptor; GPCR; integral membrane protein; antigen;
KW neuronal cell; nonhuman primate; NHP; G-protein signalling pathway;
KW apoptosis; immunogen; therapeutic; treatment; prevention; diagnostic.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH Domain 177..198
XX FT Domain /label= Transmembrane_domain_1
XX FT Domain 199..201
XX FT /label= DRP_motif
XX FT /note= Substitution of the Arg abrogates protection"
XX FT Domain 213..238
XX FT /label= Transmembrane_domain_2
XX
XX WO200022125-A2.

```

```

XX 20-APR-2000.
XX
XX 13-OCT-1999; 99WO-US21621.
XX
XX 13-OCT-1998; 98US-0104104.
XX
XX (AMHP ) AMERICAN HOME PROD CORP.
XX
XX Ozenberger BA, Kajkowski EM, Lo CF;
XX WPI: 2000-3:7982/27.
XX N-PSDB: AA252369.
XX
XX Novel G-protein-coupled receptor-like proteins and polynucleotides
XX useful for regulating apoptosis, comprises integral membrane protein
XX traversing the membrane twice -
XX
XX Example 1: Page 62-63; 68pp; English.
XX
XX The present sequence is the beta-amyloid peptide (BAP) binding protein-1
XX (BBP1). It is an integral membrane protein, that traverse the membrane
XX twice. It is related to G protein-coupled receptor (GPCR) protein
XX superfamily. It interacts with G-alpha proteins and regulates the
XX activity of G-protein signalling pathways. Bap genes are widely expressed
XX in neuronal cells of nonhuman primate (NHP) brain and overexpressed in
XX some tumours. It functions as a suppressor of apoptosis induction. BAP
XX proteins are used as immunogens to raise antibodies, useful as
XX therapeutics and as antigens in solid phase assays. They are also useful
XX as reagents to identify molecules which effect the interaction of BBP and
XX a cloned protein, that are useful in the treatment or prevention of
XX diseases associated with apoptosis. The polynucleotides are useful for
XX diagnostics.
XX
XX Note: In claim 5, the patent claims an amino acid sequence from figure 2.
XX However, figure 2 does not contain any sequence. It is inferred from the
XX disclosure that the figure 2 sequence refers to BBP1 protein, shown in
XX this sequence.
XX
XX Sequence 269 AA:
XX
XX Query Match 100.0%; Score 202; DB 21; Length 269;
XX Best Local Similarity 100.0%; Pred. No. 7e-193;
XX Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 PSGPSAPEAVTARLVGVLFVSVITGPGAVATSGAGEESLKCEDLKVGQYICKDKPKIND 60
Db 68 PSGPSAPEAVTARLVGVLFVSVITGPGAVATSGAGEESLKCEDLKVGQYICKDKPKIND 127
XX
QY 61 ATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFKPISCRNVNGYSYKVAV 120
Db 128 ATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFKPISCRNVNGYSYKVAV 187
XX
QY 121 ALSFLGWLGAADRFYLGYPALGLKFCITGFCIGSLIDFILISMQIVGPSDGSSTYIDY 180
Db 188 ALSFLGWLGAADRFYLGYPALGLKFCITGFCIGSLIDFILISMQIVGPSDGSSTYIDY 247
XX
QY 181 YGTRLTRLSITNETFRKTQLYP 202
Db 248 YGTRLTRLSITNETFRKTQLYP 269
XX
XX RESULT 3
XX AAE33877
XX ID AAE33877 standard; Protein: 269 AA.
XX AC AAE33877;
XX DT 02-MAY-2003 (first entry)
XX DE Human BBP-1 protein.
XX
XX Human; beta-amyloid peptide-binding protein; BAP; Abeta; betaAP; BBP;
XX Alzheimer's disease; AD; transgenic; transgenic animal; gene therapy;

```

KW neuroprotective; nootropic.
 XX Homo sapiens.
 OS WO200290499-A2.
 XX PN 14-NOV-2002.
 XX PD 06-MAY-2002; 2002WO-US14223.
 XX PF 09-MAY-2001; 2001US-0852100.
 XX PR (AMHP) WYETH.
 XX PA Ozenberger BA, Bard JA, Kajkowski EM, Jacobson JS, Walker SG;
 XX PI Sofia RJ, Howland DS;
 XX DR WPI; 2003-120537/11.
 XX DR N-PSDB; AAD51940.
 XX DR New human beta-amyloid peptide-binding protein, useful for diagnosing
 PT and/or treating diseases associated with aberrant expression of
 PT beta-amyloid peptide, e.g. Alzheimer's disease .
 XX Claim 4; Page 84-85; 85pp; English.
 XX The present invention relates to novel human beta-amyloid peptide (BAP;
 CC Abeta, betaAP)-binding (BBP) proteins and polynucleotides encoding such
 CC proteins. BBP sequences are useful to diagnose and/or treat diseases
 CC associated with aberrant expression of human BAP such as Alzheimer's
 CC disease (AD). They are used to generate transgenic animals. Sequences
 CC of the invention are also used in gene therapy. The present sequence
 CC is human BBP-1 protein.
 XX Sequence 269 AA;
 SQ Query Match 100.0%; Score 202; DB 24; Length 269;
 Best Local Similarity 100.0%; Pred. No. 7e-193;
 Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PSGSAPAVTARLVGVLFVSVTTGPGAVATSAGGESLKCEDLKVGQYICKDPKIND 60
 DB 68 PSGSAPAVTARLVGVLFVSVTTGPGAVATSAGGESLKCEDLKVGQYICKDPKIND 127
 QY 61 ATOEPVNCNTNTAHVSCFPAPNITCKDSSGNETHFTGNEVGFFKPISCRNVNGYSYKVA 120
 DB 128 ATOEPVNCNTNTAHVSCFPAPNITCKDSSGNETHFTGNEVGFFKPISCRNVNGYSYKVA 187
 QY 121 ALSFLGLWLGADRFYLGYPALGLLKFCTVGFPGISLIDFIDLSIMQIVGPSDSSYIIDY 180
 DB 188 ALSFLGLWLGADRFYLGYPALGLLKFCTVGFPGISLIDFIDLSIMQIVGPSDSSYIIDY 247
 QY 181 YGTRLTRLSITNETRKTQIYP 202
 DB 248 YGTRLTRLSITNETRKTQIYP 269
 RESULT 4
 AAY12358
 ID AAY12358 standard; Protein; 139 AA.
 XX AC AAY12358;
 XX DT 17-JUN-1999 (first entry)
 XX DE Human 5' EST secreted protein SEQ ID NO:389.
 XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
 KW forensic; gene therapy; chromosome mapping; signal peptide;
 KW upstream regulatory sequence; cytokine activity; cell proliferation;
 KW differentiation; haematopoiesis regulation; tissue growth regulation;
 KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; anti-inflammatory; tumour inhibition.

XX Homo sapiens.
 OS WO9906548-A2.
 XX PN 11-FEB-1999.
 XX PU 31-JUL-1998; 98WO-1801222.
 XX PF 01-AUG-1997; 97US-0905135.
 XX PK (GEST) GENSET.
 XX PA Duclert A, Duras Milne Edwards J, Lacroix B;
 XX PI WPI; 1999-153776/13.
 XX DR N-PSDB; AAX41191.
 XX DR New nucleic acids encoding human secreted proteins obtained from
 PT kidney, lung, umbilical cord, placenta and colon tissue
 PT Claim 27; Page 714-715; 824pp; English.
 XX AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for
 CC human secreted proteins, and encode the proteins given in AAY12261 to
 CC AAY12514, respectively. The proteins given represent the signal peptide
 CC and an N-terminal fragment of a secreted protein. The nucleic acid
 CC sequences can be used for producing secreted human gene products. They
 CC can also be used to develop products for diagnosis and therapy. The
 CC proteins obtained may have cytokine activity, cell
 CC proliferation/differentiation activity, haematopoiesis regulating
 CC activity, tissue growth regulating activity, reproductive hormone
 CC regulating activity, chemotactic/ chemokinetic activity, haemostatic and
 CC thrombolytic activity, receptor/ ligand activity, anti-inflammatory
 CC activity, tumour inhibition activity or other activities. The products
 CC can be used in forensic, gene therapy and chromosome mapping procedures.
 CC The sequences can also be used for obtaining corresponding promoter
 CC sequences. The nucleic acids encoding the signal peptide can be used for
 CC directing extracellular secretion of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell.
 XX Sequence' 139 AA;
 SQ Query Match 65.88; Score 133; DB 20; Length 139;
 Best Local Similarity 100.0%; Pred. No. 2.3e-124;
 Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 SGPSAPAVTARLVGVLFVSVTTGPGAVATSAGGESLKCEDLKVGQYICKDPKINDA 61
 DB 7 SGPSAPAVTARLVGVLFVSVTTGPGAVATSAGGESLKCEDLKVGQYICKDPKINDA 66
 QY 62 TOEPVNCNTNTAHVSCFPAPNITCKDSSGNETHFTGNEVGFFKPISCRNVNGYSYKVA 120
 DB 67 TOEPVNCNTNTAHVSCFPAPNITCKDSSGNETHFTGNEVGFFKPISCRNVNGYSYKVA 126
 QY 122 LSFLGLWLGADRF 134
 DB 127 LSFLGLWLGADRF 139
 RESULT 5
 AAY36021
 ID AAY36021 standard; Protein; 162 AA.
 XX AC AAY36021;
 XX DT 13-SEP-1999 (first entry)
 XX DE Extended human secreted protein sequence, SEQ ID NO. 406.
 XX Secreted protein; human; cytokine; cellular proliferation; cell movement;
 KW cellular differentiation; immune system regulator; anti-inflammatory;

KW haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
 KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
 genetic disease.
 XX Homo sapiens.
 XX W09906548-A2.
 OS 11-FEB-1999.
 PN W09931236-A2.
 XX 24-JUN-1999.
 PD 17-DEC-1998; 98WO-IB02122.
 XX 10-AUG-1998; 98US-0096115.
 PR 17-DEC-1997; 97US-0089957.
 PR 09-FEB-1998; 98US-0074121.
 PR 13-APR-1998; 98US-0081563.
 XX (GEST) GENSET.
 PA Bouqueleret L, Duclert A, Dumas Milne Edwards J, Lacroix B;
 PI WPI: 1999-385906/32.
 XX N-PSDB: AAX97705.
 DR New isolated human secreted proteins
 XX Claim 9; Page 346-347; 516pp; English.
 XX This sequence is encoded by an extended human secreted protein coding
 CC sequence of the invention. The secreted proteins can be used in treating
 CC or controlling a variety of human conditions. The secreted proteins may
 CC act as cytokines or may affect cellular proliferation or differentiation
 CC or may act as immune system regulators, haematopoiesis regulators, tissue
 CC growth regulators, regulators of reproductive hormones or cell movement
 CC or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
 CC tumour inhibition activity. The DNAs can be used in forensic procedures
 CC to identify individuals or in diagnostic procedures to identify
 CC individuals having genetic diseases resulting from abnormal expression of
 CC the genes corresponding to the extended cDNAs. They are also useful for
 CC constructing a high resolution map of the human chromosomes. They can
 CC also be used for gene therapy to control or treat genetic diseases.
 XX Sequence 162 AA;
 SQ

Query Match 55.0%; Score 111; DB 20; Length 162;
 Best Local Similarity 100.0%; Pred. No. 2.2e-102;
 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 6 APEAVTARLVGVLFVSVTTGPGAVATSGAGRESLKCEDIKVGOYICKPKINDATQEP 65
 DB 11 APEAVTARLVGVLFVSVTTGPGAVATSGAGRESLKCEDIKVGOYICKPKINDATQEP 70
 OY 66 VNCNTNYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFFKPISCRNVNGYSY 116
 DB 71 VNCNTNYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFFKPISCRNVNGYSY 121

RESULT 6
 AAY12426
 ID AAY12426 standard; Protein: 148 AA.
 AC AAY12426;
 XX 17-JUN-1999 (first entry)
 DT Human 5' EST secreted protein SEQ ID NO:457.
 XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
 KW forensic; gene therapy; chromosome mapping; signal peptide;
 KW upstream regulatory sequence; cytokine activity; cell proliferation;
 KW differentiation; haematopoiesis regulation; tissue growth regulation;
 KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; anti-inflammatory; tumour inhibition.

XX Homo sapiens.
 XX W09906548-A2.
 OS 11-FEB-1999.
 PN W09931236-A2.
 XX 24-JUN-1999.
 PD 17-DEC-1998; 98WO-IB01222.
 XX 01-AUG-1997; 97US-0905135.
 PR (GEST) GENSET.
 PA Duclert A, Dumas Milne Edwards J, Lacroix B;
 PI WPI: 1999-153778/13.
 XX N-PSDB: AAX41259.
 DR New nucleic acids encoding human secreted proteins - obtained from
 PT cDNA libraries prepared from e.g. liver, ovary, brain, prostate,
 PR kidney, lung, umbilical cord, placenta and colon tissue
 XX Claim 27; Page 763-764; 824pp; English.
 XX AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for
 CC human secreted proteins, and encode the proteins given in AAY12261 to
 CC AAY12514, respectively. The proteins given represent the signal peptide
 CC and an N-terminal fragment of a secreted protein. The nucleic acid
 CC sequences can be used for producing secreted human gene products. They
 CC can also be used to develop products for diagnosis and therapy. The
 CC proteins obtained may have cytokine activity, cell
 CC proliferation/differentiation activity, haematopoiesis regulating
 CC activity, tissue growth regulating activity, chemokinetic activity, haemostatic and
 CC thrombolytic activity, chemotactic/chemokinetic activity, anti-inflammatory
 CC activity, tumour inhibition activity or other activities. The products
 CC can be used in forensic, gene therapy and chromosome mapping procedures.
 CC The sequences can also be used for obtaining corresponding promoter:
 CC sequences. The nucleic acids encoding the signal peptide can be used for
 CC directing extracellular secretion of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell.
 XX Sequence 148 AA;
 SQ

Query Match 40.5%; Score 82; DB 20; Length 148;
 Best Local Similarity 100.0%; Pred. No. 1.7e-73;
 Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 6 APEAVTARLVGVLFVSVTTGPGAVATSGAGRESLKCEDIKVGOYICKPKINDATQEP 65
 DB 11 APEAVTARLVGVLFVSVTTGPGAVATSGAGRESLKCEDIKVGOYICKPKINDATQEP 70
 OY 66 VNCNTNYTAHVSCFPAPNITCKD 87
 DB 71 VNCNTNYTAHVSCFPAPNITCKD 92

RESULT 7
 AAU97631
 ID AAU97631 standard; Protein: 100 AA.
 AC AAU97631;
 XX 13-AUG-2002 (first entry)
 DT RNA polymerase II subunit 11 protein.
 XX RNA polymerase II subunit 11; cancer; HIV; infection;
 KW human immunodeficiency virus.
 XX Unidentified.
 OS CN1331300-A.
 PN

```
XX PD 16-JAN-2002.
XX PF 30-JUN-2000; 2000CN-0116963.
XX PR 30-JUN-2000; 2000CN-0116963.
XX PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX PI Mao Y, Xie Y;
XX PR WPI: 2002-340664/38.
XX DR N-PSDB; ABK52558.
XX SQ
PT Polypeptide-RNA polymerase II subunit I; and polynucleotide for coding
PT it.
XX Claim 1; Page 29; 32pp; Chinese.
XX CC This invention relates to the DNA and protein sequences of a novel
XX polypeptide-RNA polymerase II subunit I protein. The invention also
XX comprises a process for preparing the polypeptide of the invention by
XX DNA recombination, the application of the polypeptide in treating
XX diseases such as cancer, human immunodeficiency virus (HIV) infection,
XX etc, the antagonist of the polypeptide and its medical action, and the
XX application of the said polynucleotide are disclosed. The present
XX sequence represents the RNA polymerase II subunit II protein of the
XX invention.
XX SQ Sequence 100 AA;
Query Match 21.3%; Score 43; DB 23; Length 100;
Best Local Similarity 100.0%; Pred. No. 8.6c-15;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9 AVTARLVGVLFVSVTTGPGAVATSGAGGESLKCDLKVGQY 51
DB 14 AVTARLVGVLFVSVTTGPGAVATSGAGGESLKCDLKVGQY 56
RESULT 8
AAE33878
ID AAE33878 standard; Protein; 68 AA.
XX AC AAE33878;
XX DT 02-MAY-2003 (first entry)
XX DE Human BBP-1 protein fragment.
XX KW Human: beta-amyloid peptide-binding protein; BAP; Abeta; betaAP; BAP;
XX Alzheimer's disease; AD; transgenic; transgenic animal; gene therapy;
XX neuroprotective; nontropic.
XX OS Homo sapiens.
XX PN WO200290499-A2.
XX PR 14-NOV-2002.
XX PD
XX PF 06-MAY-2002; 2002WO-US14223.
XX PR 09-MAY-2001; 2001US-0852100.
XX PX (AMHP ) WYETH.
XX PI Ozenberger BA, Bard JA, Kajkowski EM, Jacobsen JS, Walker SG;
XX Sofia HJ, Howland DS;
XX WPI: 2003-120537/11.
XX DR N-PSDB; AAD51978.
XX PT New human: beta-amyloid peptide-binding protein, useful for diagnosing
```

```
PT and/or treating diseases associated with aberrant expression of
PT beta-amyloid peptide, e.g. Alzheimer's disease -
XX PS Example 11; Fig 9; 85pp; English.
XX CC The present invention relates to novel human beta-amyloid peptide (BAP;
XX Abeta, betaAP)-binding (BBP) proteins and polynucleotides encoding such
XX proteins. BBP sequences are useful to diagnose and/or treat diseases
XX associated with aberrant expression of human BAP such as Alzheimer's
XX disease (AD). They are used to generate transgenic animals. Sequences
XX of the invention are also used in gene therapy. The present sequence
XX is human BBP-1 protein fragment.
XX SQ Sequence 68 AA;
Query Match 8.4%; Score 17; DB 24; Length 68;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 150 GFCGIGSLIDFILISMQ 166
DB 1 GFCGIGSLIDFILISMQ 17
RESULT 9
ABH11574
ID ABH11574 standard; peptide; 225 AA.
XX AC ABH11574;
XX DT 11-JAN-2002 (first entry)
XX DE Human secreted protein homologue, SEQ ID NO:1944.
XX KW Human: cytokine; cell proliferation; cell differentiation; growth factor;
XX haematopoiesis regulation; tissue growth; immunomodulator; activin;
XX inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
XX proliferation; metastasis; cancer; tumour; haematopoietic disorder;
XX myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
XX chronic inflammatory condition; proliferative retinopathy;
XX atherosclerosis; coronary heart disease; arterial ischaemia;
XX bone disorder; osteoporosis; vascular growth disorder;
XX tissue regeneration; wound healing; infection; immune disorder;
XX cell culture; drug screening; gene therapy; antiinflammatory;
XX antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
XX cytostatic; osteopathic; vasotropic; cardiac; virucide; antibacterial;
XX antifungal; vulnerary; antitumor.
XX OS Homo sapiens.
XX PN WO200157188-A2.
XX PR 09-AUG-2001.
XX PD
XX PF 05-FEB-2001; 2001WO-US03800.
XX PR 03-FEB-2000; 2000US-0456914.
XX PR 27-APR-2000; 2000US-0560875.
XX PX (HYSE-) HYSEQ INC.
XX PI Tang Y, Liu C, Drmanac RT;
XX WPI: 2001-457740/49.
XX DR N-PSDB; ABA08818.
XX CC Human proteins and DNA encoding sequences useful for preventing,
XX treating or ameliorating a medical condition in a mammalian subject
XX e.g. arthritis and cancer -
XX Claim 20; Page 214; 1963pp; English.
XX Sequences ABH10981-ABH12330 represent 1350 novel human polypeptides, and
```

sequences ABA08225-ABA09574 represent nucleic acids encoding them. The invention also relates to vectors and recombinant host cells comprising a nucleotide of the invention, methods of producing the novel polypeptides, antibodies against the polypeptides, methods of detecting the nucleotides or polypeptides in a sample, and methods of identifying compounds which bind to polypeptides of the invention. Although novel, many of the polypeptides of the invention have homology to known proteins, thereby giving an insight into their probable biological activities, and hence potential therapeutic applications. The polypeptides of the invention may have various activities, including cytokine, cell proliferation or cell differentiation activities; stem cell growth factor activity; haematopoietic regulatory activity; tissue growth activity; immunomodulatory activity; activin- or inhibin-related activities; chemotactic or chemokinetic activities; haemostatic, thrombotic or thrombolytic activities; receptor or ligand activities; or may be involved in oncogenesis, cancer cell proliferation or metastasis. Depending on their biological activities, polypeptides and nucleotides of the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound healing (e.g., of burns, incisions and ulcers), while those with immunomodulatory activities may be used in the treatment of viral, bacterial and fungal infections in addition to immune disorders. Polypeptides with growth factor activity may be used in cell cultures to promote cell growth. For example, such polypeptides may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease or accidental damage. The polypeptides and nucleotides may also be used in the diagnosis of the above conditions, and in drug screening techniques. The present sequence represents a novel human polypeptide of the invention.

XX Sequence 225 AA;

Query Match 4.5%; Score 9; DB 22; Length 225;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 LGADRFYLG 137
|||||||
DB 176 LGADRFYLG 184

RESULT 10
ARB77688
ID ARB77688 standard; peptide; 24 AA.

XX AC ARB77688;

XX DT 01-JUL-2002 (first entry)

XX DE New peptide vector#4.

XX KW Intracellular delivery; transfection agent; cancer; infectious disease; peptide vector.

XX OS Synthetic.

XX PN WO200210201-A2.

XX PD 07-FEB-2002.

XX PF 26-JUL-2001; 2001WO-US23406.

XX PR 31-JUL-2000; 2000US-221932P.

XX PA (ACTI-) ACTIVE MOTIF.

XX PA (CNRS) CENT NAT RECH SCI.

XX 21 Divida G, Morris M, Mery J, Heitz F, Fernandez J, Archdeacon J;
PI Horndorff K;
XX
XX WPI; 2002-329441/36.

XX Transfection agent that comprises a peptide comprising hydrophobic and
PT hydrophilic domain and having amino acid residues of specified length
PT is useful for a non-covalent association with and transport of a
XX heterologous compound into a cell -
XX Example 2; Page 61; 156pp; English.

XX The invention relates to a transfection agent comprising a peptide of
CC about 16-30 amino acids in length. Peptides of the invention comprise
CC a hydrophobic domain, a hydrophilic domain, optionally a spacer sequence
CC between the domains and a functional group conjugated to at least one
CC terminal of the peptide. Peptides of the invention are useful for a
CC non-covalent association with and transport of a heterologous compound
CC into a cell. They are also useful for promoting the cellular
CC internalisation of at least one member e.g. peptide, proteins,
CC antibodies, their derivatives and/or conjugates. They may form part of a
CC pharmaceutical composition to deliver the compound selected from a
CC diagnostic or therapeutic compound, to treat at least one condition: such
CC as cancer or an infectious disease, or which targets a cancerous cell or
CC pathogen-infected cell and to deliver a peptide or inhibitor that
CC disrupts the activity of the enzyme. The agent of the invention has a
CC transfection efficiency of at least 5% for at least two of the members of
CC the group of the compounds. The agent has a good delivery efficiency for
CC a broad spectrum of compounds and cell types, has a low toxicity, are
CC easy to handle and easy to formulate in conjunction with the many
CC different compound types that it can deliver. The peptides are serum
CC sensitive, thus they code particularly well for systemic and/or localised
CC in patients. The current sequence represents a new amphipathic peptide
CC vector of the invention that contains a cationic nuclear localisation
CC sequence.

XX Sequence 24 AA;

Query Match 4.0%; Score 8; DB 23; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 LFLGWLGA 131
|||||||
DB 3 LFLGWLGA 10

RESULT 11
ARB81176
ID ARB81176 standard; peptide; 27 AA.

XX AC ARB81176;

XX DT 25-NOV-2002 (first entry)

XX DE Signal sequence based peptide 1 MPS peptide.

XX KW Red blood cell; RBC; virus; vertebrate; therapeutic; tissue imaging;
KW qp41; membrane translocation.

XX OS Human immunodeficiency virus type 1.

XX OS Rhesus macaque polyoma virus.

XX PN WO200260416-A1.

XX PD 08-AUG-2002.

XX PF 01-FEB-2002; 2002WO-GB00437.

XX PR 01-FEB-2001; 2001GB-0002561.

XX PR 16-FEB-2001; 2001US-269528P.

PA (GEND-) GENDEL LTD.
XX
XX PI Mchale AP, Craig R;
XX DR WPI: 2002-643355/69.
XX
XX PT Delivering agent to target site in vertebrate comprises loading red
XX PT blood cell with virus or virus-like particle comprising agent,
XX PT sensitizing cell, introducing cell into vertebrate and applying energy
XX PT to release virus particle from cell.
XX PS Disclosure: Page 55: 87pp; English.
XX
XX CC The invention relates to delivering an agent to a target site in a
XX CC vertebrate. The method involves (a) loading a red blood cell (RBC) with a
XX CC virus or a virus-like particle (i) comprising an agent; (b) sensitizing
XX CC RBC to render it more susceptible to disruption than unsensitized RBC;
XX CC (c) introducing RBC into a vertebrate, and (d) applying energy to release
XX CC (i). Steps (a) and (b) may be performed in any order. (ii) (RBC loaded
XX CC with a virus or a virus-like particle comprising a therapeutic agent) is
XX CC useful for the delivery of a therapeutic agent to a target site in a
XX CC vertebrate, or in the preparation of a medicament for delivery of a
XX CC therapeutic agent to a target site in a vertebrate. (ii) is also useful
XX CC for delivering one or more agents to a vertebrate and for treating or
XX CC preventing a disease. The method is useful for delivering agents such as
XX CC those useful for imaging of tissues in vivo or ex vivo, preferably for
XX CC delivering an agent to a subcellular organelle such as nucleus,
XX CC mitochondria, Golgi or endoplasmic reticulum. The present sequence
XX CC represents a signal sequence based peptide: 1. MPS peptide, a chimera of
XX CC the hydrophobic terminal domain of viral gp41 protein and the nuclear
XX CC localisation signal (NLS) from the SV40 large antigen. This fragment has
XX CC been found to be active in membrane translocation.
XX
XX SQ Sequence 27 AA;
Query Match 4.0%; Score 8; DB 23; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 124 LFLGLWLGA 131
DB 11111111
3 LFLGLWLGA 10
RESULT 12
ABG78990
ID ABG78990 standard; Peptide: 27 AA.
XX AC ABG78990;
XX DT 15-NOV-2002 (first entry)
XX DE Cell penetrating peptide signal-peptide II.
XX KW Cell penetrating peptide; cancer; tumour; melanoma; thymoma;
XX KW lymphoma; sarcoma; lung cancer; non-Hodgkin's lymphoma; leukaemia;
XX KW Hodgkin's lymphoma; uterine cancer; cervical cancer; bladder cancer;
XX KW kidney cancer; adenocarcinoma; breast cancer; prostate cancer;
XX KW ovarian cancer; pancreatic cancer; vaccine; dendritic cell;
XX KW tumour infiltrating lymphocyte; TIL; human leukocyte antigen; HLA;
XX KW cytostatic.
XX OS Unidentified.
XX PN WO200264057-A2.
XX PF 22-AUG-2002.
XX PD 15-FEB-2002; 2002WO-US05212.
XX PF 15-FEB-2001; 2001US-265687P.
XX PR (BAYU) BAYLOR COLLEGE MEDICINE.
XX PA

XX Wang R;
XX WPI: 2002-627577/67.
XX
XX PT Novel composition for treating a disease in an animal, comprises an
XX PT immune effector cell and cell penetrating peptide associated with an
XX PT antigen or antibody.
XX FS Disclosure: Page 11: 61pp; English.
XX
XX CC The invention relates to a composition (i) comprising an immune effector
XX CC cell and a cell penetrating peptide (CPP) associated with an antigen or
XX CC antibody. Also included are (ii) a vaccine comprising (i), CPP
XX CC associated with an antigen, and a pharmaceutically acceptable carrier
XX CC and (2) preparing a composition for a disease, by providing (i)
XX CC and CPP associated with an antigen for disease, and introducing the
XX CC antigen-associated CPP to (i), where antigen enters into the cell.
XX CC The antigens are, for example, tumour antigen derived epitopes
XX CC recognised by tumour infiltrating lymphocytes (TIL) of HLA (human
XX CC leukocyte antigen) class I or II. The composition is useful for enhancing
XX CC immunity in an animal to a disease, by administering a mature dendritic
XX CC cell comprising CPP associated with an antigen to disease, to the animal,
XX CC such that following the administration, animal is protected from disease,
XX CC where the animal comprises both CD4+ and CD8+ T cells. It is also useful
XX CC for treating a disease (e.g. cancer, tumour, melanoma, thymoma,
XX CC lymphoma, sarcoma, lung cancer, non-Hodgkin's lymphoma, leukaemia,
XX CC Hodgkin's lymphoma, uterine cancer, cervical cancer, bladder cancer,
XX CC kidney cancer, adenocarcinoma, breast cancer, prostate cancer,
XX CC ovarian cancer and pancreatic cancer). The animal is further subjected to
XX CC a cancer treatment including surgery, radiation, chemotherapy or gene
XX CC therapy. The administration of (i), preferably dendritic cell is prior
XX CC to, subsequent to or concurrent with, the cancer treatment. The present
XX CC sequence is cell penetrating peptide of the invention.
XX
XX SQ Sequence 27 AA;
Query Match 4.0%; Score 8; DB 23; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 124 LFLGLWLGA 131
DB 11111111
3 LFLGLWLGA 10
RESULT 13
AAE23685
ID AAE23685 standard; peptide: 27 AA.
XX AC AAE23685;
XX DT 10-SEP-2002 (first entry)
XX DE Fluorescently labelled RNA binding peptide #2.
XX KW RNA binding protein; mRNA quantification; gene expression.
XX OS Unidentified.
XX PN WO200227031-A2.
XX PF 04-APR-2002.
XX PF 28-SEP-2001; 2001WO-US30438.
XX PR 28-SEP-2000; 2000US-236407P.
XX PA (CELL-) CELLOMICS INC.
XX PI Busa WB;
XX WPI: 2002-452280/48.
XX DR

XX
 PT Quantifying target gene expression in living cells that possess a
 PT target gene of interest tagged with the binding site for an RNA binding
 PT protein and fluorescently labeled RNA binding polypeptide including a:
 PT RNA binding domain -
 XX
 PS Claim 45; Page 44; 51pp; English.
 XX
 CC The present invention relates to a method of quantifying the expression
 CC of target genes in living cells. The method involves providing cells that
 CC possess a target gene of interest which has been tagged with the binding
 CC site for an RNA binding protein and a fluorescently labelled RNA binding
 CC polypeptide that includes an RNA binding domain and calculating the
 CC quantity of target gene expression in the cells using fluorescence
 CC signalling techniques. The method is useful for quantifying expression
 CC of one or more target genes in living cells which comprise two or more
 CC distinct populations of cells. It is used to quantitate the expression
 CC of any target gene, including expression of protein-encoding messenger
 CC RNA genes, ribosomal RNA encoding genes and transfer RNA encoding genes
 CC so long as the RNA expression product from the target gene possesses a
 CC sequence or structure (the RNA tag) that is bound specifically by the
 CC RNA binding polypeptide being used. The present sequence is a
 CC fluorescently labelled RNA binding peptide.
 XX
 SQ Sequence 27 AA;
 Query Match 4.0%; Score 8; DB 23; Length 27;
 Best Local Similarity 100.0%; Pred. No. 1.9; Mismatches 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0;
 QY 124 LFLGWLGA 131
 DB 3 LFLGWLGA 10
 I:|||||
 3 LFLGWLGA 10
 RESULT 14
 ASB77687
 ID ASB77687 standard; peptide: 27 AA.
 XX
 AC ASB77687;
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE New peptide vector#3.
 XX
 KW Intracellular delivery: transfection agent; cancer; infectious disease;
 KW peptide vector.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 7
 FT /note= "residue may be substituted with Phe"
 FT Misc-difference 23
 FT /note= "residue may be substituted with Ser"
 XX
 PN WO200210201-A2.
 XX
 PD 07-FEB-2002.
 XX
 XX 26-JUL-2001; 2001WO-US23406.
 XX
 XX 31-JUL-2000; 2000US-221932P.
 XX
 PA (ACTI-) ACTIVE MOTIF.
 PA (CNRS) CENT NAT RECH SCI.
 XX
 XX Divida G, Morris M, Mery J, Heitz F, Fernandez C, Archdeacon C;
 PI Horndorp K;
 XX
 XX WPI; 2002-329441/36.
 DR
 XX Transfection agent that comprises a peptide comprising hydrophobic and

PT hydrophilic domain and having amino acid residues of specified length
 PT is useful for a non-covalent association with and transport of a
 XX heterologous compound into a cell -
 PS Example 2; Page 61; 156pp; English.
 XX
 CC The invention relates to a transfection agent comprises a peptide of
 CC about 16-30 amino acids in length. Peptides of the invention comprise
 CC a hydrophobic domain, a hydrophilic domain, optionally a spacer sequence
 CC between the domains and a functional group conjugated to at least one
 CC terminal of the peptide. Peptides of the invention are useful for a
 CC non-covalent association with and transport of a heterologous compound
 CC into a cell. They are also useful for promoting the cellular
 CC internalisation of at least one member e.g. peptide, proteins,
 CC antibodies, their derivatives and/or conjugates. They may form part of a
 CC pharmaceutical composition to deliver the compound selected from a
 CC diagnostic or therapeutic compound, to treat at least one condition such
 CC as cancer or an infectious disease, or which targets a cancerous cell or
 CC pathogen-infected cell and to deliver a peptide or inhibitor that
 CC disrupts the activity of the enzyme. The agent of the invention has a
 CC transfection efficiency of at least 5% for at least two of the members of
 CC the group of the compounds. The agent has a good delivery efficiency for
 CC a broad spectrum of compounds and cell types, has a low toxicity, are
 CC easy to handle and easy to formulate in conjunction with the many
 CC different compound types that it can deliver. The peptides are serum
 CC sensitive, thus they bode particularly well for systemic and/or localised
 CC in patients. The current sequence represents a new amphipathic peptide
 CC vector of the invention that contains a cationic nuclear localisation
 CC sequence.
 XX
 SQ Sequence 27 AA;
 Query Match 4.0%; Score 8; DB 23; Length 27;
 Best Local Similarity 100.0%; Pred. No. 1.9; Mismatches 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0;
 QY 124 LFLGWLGA 131
 DB 3 LFLGWLGA 10
 I:|||||
 3 LFLGWLGA 10
 RESULT 15
 AAU78348
 ID AAU78348 standard; Peptide: 27 AA.
 XX
 AC AAU78348;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Signal sequence based peptide I.
 XX
 KW Membrane translocation signal; signal sequence based peptide I;
 KW red blood cell vehicle; polypeptide delivery; viral gp41 protein;
 KW simian virus 40; SV40; large antigen.
 XX
 CS Rhesus macaque polyoma virus.
 CS Synthetic.
 XX
 FN WO200207752-A2.
 XX
 PD 31-JAN-2002.
 XX
 XX 24-JUL-2001; 2001WO-GH03327.
 XX
 XX 24-JUL-2000; 2000WO-GH02848.
 PR 29-AUG-2000; 2000WO-GS03056.
 PR 01-FEB-2001; 2001WO-GS00417.
 PR 16-FEB-2001; 2001US-0785802.
 XX
 XX (GENO-) GENDEL LTD.
 XX
 XX Craig K;
 PI

DR WPI: 2002-280593/32.
XX
PT Preparing a red blood cell vehicle suitable for delivering an agent to
PT a target site in a vertebrate due to loading the red blood cell with an
PT agent-membrane translocation sequence . XX
XX
PS Disclosure: Page 44; 135pp; English.
PS
XX The invention describes a method of preparing a red blood cell vehicle
CC suitable for delivering an agent to a target site in a vertebrate
CC comprising providing a red blood cell and loading the red blood cell with
CC an agent-MTS (membrane translocation sequence) conjugate. The red blood
CC cells produced may be used in the preparation of a medicament for
CC delivery of an agent to or at a target site and of one or more agents to
CC a vertebrate. The agent is actively released from the red blood cell
CC vehicle by application of a stimulus to disrupt the red blood cell
CC vehicle. This sequence represents signal sequence based peptide 1, a
CC chimeric peptide of the hydrophobic terminal domain of the viral gp41
CC protein and the nuclear localisation signal from simian virus 40
CC (SV40) large antigen, one of the membrane translocation peptides tested
CC in the invention.
XX
SQ Sequence 27 AA:

Query Match 4.0%; Score 8; DB 23; Length 27;
Best Local Similarity 100.0%; Pred. No. 1 9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 LFLGWLGA 131
11111111
Dh 3 LFLGWLGA 10

Search completed: September 26, 2003, 17:39:24
Job time : 103.289 secs

GenCore version: 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:36:55 ; Search time 36.1021 seconds
(without alignments)
230.739 Million coll. updates/sec

Title: US-09-852-100A-2_COPY_68_269

Perfect score: 202
Sequence: 1 PGGSPAPAVTARLVGVLF.....TRITRSTTNETFKTKTLYP 202

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310856 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA:

- 1: /cgn2.6/ptodata/1/1aa/5A.COMB.pep.*
- 2: /cgn2.6/ptodata/1/1aa/5B.COMB.pep.*
- 3: /cgn2.6/ptodata/1/1aa/6A.COMB.pep.*
- 4: /cgn2.6/ptodata/1/1aa/6B.COMB.pep.*
- 5: /cgn2.6/ptodata/1/1aa/PCUS.COMB.pep.*
- 6: /cgn2.6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	4.0	358	US-09-411-628-11	Sequence 11, Appl
2	7	3.5	114	US-09-222-938A-13	Sequence 13, Appl
3	7	3.5	183	US-09-252-991A-21850	Sequence 21850, A
4	7	3.5	204	US-09-134-001C-5239	Sequence 5239, Ap
5	7	3.5	213	US-09-163-445B-1	Sequence 1, Appl
6	7	3.5	245	US-09-393-634-55	Sequence 55, Appl
7	7	3.5	315	US-09-393-634-56	Sequence 56, Appl
8	7	3.5	451	US-09-345-4735-45	Sequence 45, Appl
9	7	3.5	462	US-09-328-352-6888	Sequence 6888, Ap
10	7	3.5	483	US-09-352-991A-29267	Sequence 29267, A
11	7	3.5	617	US-09-345-4735-44	Sequence 44, Appl
12	7	3.5	643	US-09-252-991A-26617	Sequence 26617, A
13	7	3.5	673	US-09-063-950-2	Sequence 2, Appl
14	7	3.5	673	US-09-996-243-52	Sequence 52, Appl
15	7	3.5	707	US-09-252-991A-29401	Sequence 29401, A
16	7	3.5	2519	US-09-413-814-42	Sequence 42, Appl
17	6	3.0	23	US-08-176-508-121	Sequence 121, App
18	6	3.0	23	US-08-471-052A-121	Sequence 121, App
19	6	3.0	23	US-08-189-331-121	Sequence 121, App
20	6	3.0	23	US-08-471-939-121	Sequence 121, App
21	6	3.0	23	US-08-471-800-121	Sequence 121, App
22	6	3.0	23	US-08-471-068-121	Sequence 121, App
23	6	3.0	38	US-09-461-325-284	Sequence 284, App
24	6	3.0	46	US-09-489-847-131	Sequence 131, App
25	6	3.0	61	US-08-266-321-3	Sequence 3, Appl
26	6	3.0	61	US-08-467-527A-3	Sequence 3, Appl
27	6	3.0	61	US-08-467-528-3	Sequence 3, Appl

```

28 6 3.0 70 4 US-09-205-258-298 Sequence 258, App
29 6 3.0 72 3 US-08-905-223-314 Sequence 314, App
30 6 3.0 101 4 US-09-489-847-306 Sequence 306, App
31 6 3.0 117 4 US-09-228-986-111 Sequence 111, App
32 6 3.0 120 4 US-08-728-742A-1 Sequence 1, Appli
33 6 3.0 123 2 US-08-675-508-2 Sequence 2, Appli
34 6 3.0 123 3 US-09-203-939-2 Sequence 2, Appli
35 6 3.0 123 3 US-09-203-939-6 Sequence 6, Appli
36 6 3.0 123 3 US-09-251-835-2 Sequence 2, Appli
37 6 3.0 123 3 US-09-251-835-6 Sequence 6, Appli
38 6 3.0 123 3 US-09-318-503-2 Sequence 2, Appli
39 6 3.0 123 3 US-09-318-503-6 Sequence 6, Appli
40 6 3.0 123 3 US-09-038-261A-2 Sequence 2, Appli
41 6 3.0 123 3 US-09-038-261A-6 Sequence 6, Appli
42 6 3.0 123 4 US-09-564-329A-6 Sequence 6, Appli
43 6 3.0 123 4 US-09-564-329A-6 Sequence 6, Appli
44 6 3.0 146 1 US-08-688-609-6 Sequence 6, Appli
45 6 3.0 146 3 US-09-002-832-6 Sequence 6, Appli

```

ALIGNMENTS

```

RESULT 1
US-09-411-628-11
: Sequence 11, Application US/09411628
: Patent No. 6428994
: GENERAL INFORMATION:
: APPLICANT: University of Southern California
: TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
: FILE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
: FILE REFERENCE: 13761-707
: CURRENT APPLICATION NUMBER: US/09/411,628
: EARLIER FILING DATE: 1999-10-01
: EARLIER APPLICATION NUMBER: US 60/102,906
: EARLIER FILING DATE: 1998-10-02
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 11
: LENGTH: 358
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-411-628-11

```

```

Query Match 4.0%; Score 8; DB 4; Length 358;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

CY 138 YPALGLLK 145
DB 259 YPALGLLK 266

```

```

RESULT 2
US-09-222-938A-13
: Sequence 13, Application US/09222938A
: Patent No. 6437108
: GENERAL INFORMATION:
: APPLICANT: Youngman, Philip
: APPLICANT: Fritz, Chrisia
: APPLICANT: Murphy, Christopher
: APPLICANT: Guzman, Luz-Maria
: TITLE OF INVENTION: ESSENTIAL BACTERIAL GENES AND THEIR USE
: FILE REFERENCE: 07334/060001
: CURRENT APPLICATION NUMBER: US/09/222,938A
: CURRENT FILING DATE: 1998-12-30
: NUMBER OF SEQ ID NOS: 102
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 13
: LENGTH: 114
: TYPE: PRT
: ORGANISM: Streptococcus pneumoniae
US-09-222-938A-13

```

```
Query Match      3.5%: Score 7; DB 4; Length 114;
Best Local Similarity 100.0%: Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
Db 69 AVATSAG 75

RESULT 3
US-09-252-991A-21850
; Sequence 5239, Application US/09252991A
; Patent No. 6561795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21850
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21850

Query Match      3.5%: Score 7; DB 4; Length 183;
Best Local Similarity 100.0%: Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSGPSAP 7
Db 131 PSGPSAP 137

RESULT 4
US-09-134-001C-5239
; Sequence 5239, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134.001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR FILING DATE: 1997-11-08
; PRIOR FILING DATE: 1997-11-08
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5239
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5239

Query Match      3.5%: Score 7; DB 4; Length 204;
Best Local Similarity 100.0%: Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
Db 115 AVATSAG 121

RESULT 5
```

```
US-09-163-445B-1
; Sequence 1, Application US/09163445B
; Patent No. 6472377
; GENERAL INFORMATION:
; APPLICANT: Fritz, Christian
; APPLICANT: Youngman, Philip
; TITLE OF INVENTION: USE OF YNES, ESSENTIAL BACTERIAL GENES AND POLYPEPTIDES
; FILE REFERENCE: 06286-C90001
; CURRENT APPLICATION NUMBER: US/09/163.445B
; CURRENT FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: US 60/070,116
; PRIOR FILING DATE: 1997-12-31
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-163-445B-1

Query Match      3.5%: Score 7; DB 4; Length 213;
Best Local Similarity 100.0%: Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
Db 105 AVATSAG 111

RESULT 6
US-09-393-634-55
; Sequence 55, Application US/09393634
; Patent No. 6558910
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Ryba, Nick
; APPLICANT: Mueller, Ken
; APPLICANT: Hoon, Mark
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: SP, a No. 6558910el Family of Taste Receptors
; FILE REFERENCE: 02307E-098000US
; CURRENT APPLICATION NUMBER: US/09/393.634
; CURRENT FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 55
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human GR11
; NAME/KEY: MOD_RES
; LOCATION: (1)..(245)
; OTHER INFORMATION: Xaa any amino acid
US-09-393-634-55

Query Match      3.5%: Score 7; DB 4; Length 245;
Best Local Similarity 100.0%: Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 SLIDFIL 162
Db 42 SLIDFIL 43

RESULT 7
US-09-393-634-56
; Sequence 56, Application US/09393634
; Patent No. 6558910
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Zuker, Charles S.
/ APPLICANT: Adler, Jon Elliot
/ APPLICANT: Ryba, Nick
/ APPLICANT: Muccler, Ken
/ APPLICANT: Hoon, Mark
/ APPLICANT: The Regents of the University of California
/ APPLICANT: The Government of the United States of America
/ APPLICANT: as represented by the Secretary of the
/ APPLICANT: Department of Health and Human Services
/ TITLE OF INVENTION: SF, a No. 6558903el Family of Taste Receptors
/ FILE REFERENCE: C2307E-C98003US
/ CURRENT APPLICATION NUMBER: US/09/333,634
/ CURRENT FILING DATE: 1999-09-10
/ NUMBER OF SEQ ID NOS: 92
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 56
/ LENGTH: 315
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: MOD_RES
/ LOCATION: (1)..(315)
/ OTHER INFORMATION: Xaa - any amino acid
US-09-393-634-56

Query Match 3.5%; Score 7; DB 4; Length 315;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 156 SLIDFIL 162
Db 42 SLIDFIL 48

RESULT 8
US-09-345-473E-45
/ Sequence 45, Application US/09345473E
/ Patent No. 6558903
/ GENERAL INFORMATION:
/ APPLICANT: Hodge, Martin
/ TITLE OF INVENTION: No. 6558903el Kinases and Uses Thereof
/ FILE REFERENCE: 35800/183781
/ CURRENT APPLICATION NUMBER: US/09/345,473E
/ CURRENT FILING DATE: 1999-06-30
/ NUMBER OF SEQ ID NOS: 62
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 45
/ LENGTH: 451
/ TYPE: PRT
/ ORGANISM: Mus musculus
US-09-345-473E-45

Query Match 3.5%; Score 7; DB 4; Length 451;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 ALSFLG 127
Db 416 ALSFLG 422

RESULT 9
US-09-328-352-6888
/ Sequence 6888, Application US/09328352
/ Patent No. 6562958
/ GENERAL INFORMATION:
/ APPLICANT: Gary L. Breton et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ADINETOBACTER
/ FILE REFERENCE: GTC99-03PA
/ CURRENT APPLICATION NUMBER: US/09/328,352

/ GENERAL INFORMATION:
/ APPLICANT: Hodge, Martin
/ TITLE OF INVENTION: No. 6558903el Kinases and Uses Thereof
/ FILE REFERENCE: 35800/183781
/ CURRENT APPLICATION NUMBER: US/09/345,473E
/ CURRENT FILING DATE: 1999-06-30
/ NUMBER OF SEQ ID NOS: 62
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 44
/ LENGTH: 617
/ TYPE: PRT
/ ORGANISM: Mus musculus
US-09-345-473E-44

Query Match 3.5%; Score 7; DB 4; Length 617;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 ALSFLG 127
Db 562 ALSFLG 588

/ GENERAL INFORMATION:
/ APPLICANT: Zuker, Charles S.
/ APPLICANT: Adler, Jon Elliot
/ APPLICANT: Ryba, Nick
/ APPLICANT: Muccler, Ken
/ APPLICANT: Hoon, Mark
/ APPLICANT: The Regents of the University of California
/ APPLICANT: The Government of the United States of America
/ APPLICANT: as represented by the Secretary of the
/ APPLICANT: Department of Health and Human Services
/ TITLE OF INVENTION: SF, a No. 6558903el Family of Taste Receptors
/ FILE REFERENCE: C2307E-C98003US
/ CURRENT APPLICATION NUMBER: US/09/333,634
/ CURRENT FILING DATE: 1999-09-10
/ NUMBER OF SEQ ID NOS: 92
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 56
/ LENGTH: 315
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: MOD_RES
/ LOCATION: (1)..(315)
/ OTHER INFORMATION: Xaa - any amino acid
US-09-393-634-56

Query Match 3.5%; Score 7; DB 4; Length 462;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 ALSFLG 127
Db 334 ALSFLG 340

RESULT 10
US-09-252-991A-29267
/ Sequence 29267, Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 62/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 29267
/ LENGTH: 483
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29267

Query Match 3.5%; Score 7; DB 4; Length 483;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 139 PALGLK 145
Db 438 PALGLK 444

RESULT 11
US-09-345-473E-44
/ Sequence 44, Application US/09345473E
/ Patent No. 6558903
/ GENERAL INFORMATION:
/ APPLICANT: Hodge, Martin
/ TITLE OF INVENTION: No. 6558903el Kinases and Uses Thereof
/ FILE REFERENCE: 35800/183781
/ CURRENT APPLICATION NUMBER: US/09/345,473E
/ CURRENT FILING DATE: 1999-06-30
/ NUMBER OF SEQ ID NOS: 62
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 44
/ LENGTH: 617
/ TYPE: PRT
/ ORGANISM: Mus musculus
US-09-345-473E-44

Query Match 3.5%; Score 7; DB 4; Length 617;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 ALSFLG 127
Db 562 ALSFLG 588
```

RESULT 12

US-09-252-991A-26617
Sequence 26617, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 26617
LENGTH: 643
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26617

Query Match 3.5%, Score 7; DB 4; Length 643;

Best Local Similarity 100.0%; Pred. No. 92;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GPSAPEA 9

|||||

DB 1 GPSAPEA 7

RESULT 13

US-09-063-950-2
Sequence 2, Application US/09063950C
Patent No. 6225085
GENERAL INFORMATION:
APPLICANT: Holtzman, Douglas A.
TITLE OF INVENTION: NOVEL LRSG PROTEIN AND NUCLEIC ACID MOLECULES AND USES
FILE REFERENCE: MEI-019
CURRENT APPLICATION NUMBER: US/09/063.950C
CURRENT FILING DATE: 1998-04-21
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 673
TYPE: PRT
ORGANISM: Homo sapiens
US-09-063-950-2

Query Match 3.5%, Score 7; DB 3; Length 673;

Best Local Similarity 100.0%; Pred. No. 96;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 ITNETFR 196

|||||

DB 115 ITNETFR 121

RESULT 14

US-09-996-243-52
Sequence 52, Application US/09996243
Patent No. 6478825
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Eric
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman

APPLICANT: Gerber, Hauspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin I.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William L.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
Acids Encoding the Same
FILE REFERENCE: P2730P1C13
CURRENT APPLICATION NUMBER: US/09/996.243
CURRENT FILING DATE: 2002-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217

; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088875
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088905
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24

; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 3.5%; Score 7; DB 4; Length 673;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 :TNETER 1%
DB 115 ITNETER 121

RESULTS 15

US-59-252-99A-29401
; Sequence 29401, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29401
; LENGTH: 707
; TYPE: PRT

: ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29401

Query Match	3.5%	Score 7:	DB 4:	Length 707:
Best Local Similarity	100.0%	Pred. No. 1e+02:	0:	Indels
Matches	7:	Conservative	0:	Mismatches
			0:	Gaps
QY	100	VGFEKPI	106	
		{		
Db	36	VGFEKPI	42	

Search completed: September 26, 2003, 17:43:23
Job time : 37.1021 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

Title: US-09-852-100A-2_COPY_68_259
 Perfect score: 202
 Sequence: 1 PSGSFAPEAVTARLVGVLPF.....TRJRLSITNETKTKQLNP 202

Scoring table: OLIGO 60.0 , Gapext 50.0

Searched: 566894 seqs, 151307093 residues

```

Word size : 0
Total number of hits satisfying chosen parameters: 566804

```

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Listing first 45 summaries

```

Database : Published_Applications_RA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp:*
2: /cgn2_6/ptodata/2/pubpaa/ECT_NEW_PUB.pcp:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp:*
6: /cgn2_6/ptodata/2/pubpaa/ECTUS_PUBCOMB.pcp:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query ³		Length	DB	ID	Description
		Match	Length				
1	202	100.0	269	9	US-09-852-100A-2		Sequence 2, Ap
2	202	100.0	269	10	US-09-833-503A-2		Sequence 2, Ap
3	202	100.0	269	15	US-10-139-88-2		Sequence 2, Ap
4	8	4.0	21	15	US-10-226-956-288		Sequence 288,
5	8	4.0	21	15	US-10-311-088-306		Sequence 306,
6	8	4.0	24	11	US-09-915-914B-28		Sequence 28, A
7	8	4.0	27	10	US-09-795-802A-8		Sequence 8, Ap
8	8	4.0	27	11	US-09-965-876A-2		Sequence 2, Ap
9	8	4.0	27	12	US-09-315-914B-27		Sequence 27, A
10	8	4.0	27	15	US-10-077-555-8		Sequence 8, Ap
11	8	4.0	221	10	US-09-833-503A-6		Sequence 6, Ap
12	8	4.0	221	11	US-09-992-560A-82		Sequence 82, A
13	8	4.0	221	11	US-09-524-740-62		Sequence 82, A
14	8	4.0	221	11	US-09-974-879-230		Sequence 230,
15	8	4.0	221	12	US-09-952-095B-82		Sequence 82, A

16	8	4.0	221	12	US-10-154-678-82	Sequence 82, Appl
17	8	4.0	221	12	US-09-999-570-82	Sequence 82, Appl
18	8	4.0	221	15	US-10-000-489-82	Sequence 82, Appl
19	8	4.0	221	15	US-10-000-986-82	Sequence 82, Appl
20	8	4.0	221	15	US-10-199-881-6	Sequence 6, Appl
21	8	4.0	222	11	US-09-305-736-230	Sequence 230, App
22	8	4.0	230	11	US-09-796-753-50	Sequence 50, Appl
23	8	4.0	247	11	US-09-796-753-48	Sequence 48, Appl
24	8	4.0	358	12	US-10-174-794-11	Sequence 11, Appl
25	7	3.5	114	15	US-10-154-251-13	Sequence 13, Appl
26	7	3.5	122	9	US-09-867-550-80	Sequence 80, Appl
27	7	3.5	173	11	US-09-946-374-332	Sequence 332, App
28	7	3.5	173	12	US-10-015-387A-332	Sequence 332, App
29	7	3.5	173	12	US-10-006-130A-332	Sequence 332, App
30	7	3.5	173	12	US-10-006-172A-332	Sequence 332, App
31	7	3.5	173	12	US-10-015-392A-332	Sequence 332, App
32	7	3.5	173	12	US-10-017-253A-332	Sequence 332, App
33	7	3.5	173	12	US-10-017-506A-332	Sequence 332, App
34	7	3.5	173	15	US-10-006-856A-332	Sequence 332, App
35	7	3.5	173	15	US-10-006-818A-332	Sequence 332, App
36	7	3.5	173	15	US-10-015-353A-332	Sequence 332, App
37	7	3.5	173	15	US-10-015-869A-332	Sequence 332, App
38	7	3.5	173	15	US-10-012-121A-332	Sequence 332, App
39	7	3.5	173	15	US-10-006-116A-332	Sequence 332, App
40	7	3.5	173	15	US-10-006-117A-332	Sequence 332, App
41	7	3.5	173	15	US-10-017-527A-332	Sequence 332, App
42	7	3.5	173	15	US-10-013-913A-332	Sequence 332, App
43	7	3.5	173	15	US-10-007-194A-332	Sequence 332, App
44	7	3.5	173	15	US-10-013-432A-332	Sequence 332, App
45	7	3.5	173	15	US-10-011-671A-332	Sequence 332, App

ALIGNMENTS

```

1  RESULT 1
2  US-09-852-100A-2
3  ; Sequence 2, Application US/09852-100A
4  ; Patent No. US20020056267A1
5  ; GENERAL INFORMATION:
6  ; APPLICANT: American Home Products
7  ; TITLE OF INVENTION: Beta-amylid Peptide-B
8  ; TITLE OF INVENTION: Same
9  ; FILE REFERENCE: AHP981261p2
10 ; CURRENT APPLICATION NUMBER: US/09/852,100A
11 ; CURRENT FILING DATE: 2001-05-09
12 ; PRIOR APPLICATION NUMBER: US 09/172,590
13 ; PRIOR FILING DATE: 1998-10-14
14 ; PRIOR APPLICATION NUMBER: US 60/104,104
15 ; PRIOR FILING DATE: 1998-10-13
16 ; PRIOR APPLICATION NUMBER: PTC/US99/21621
17 ; PRIOR FILING DATE: 1999-10-13
18 ; PRIOR APPLICATION NUMBER: US 09/060,609
19 ; PRIOR FILING DATE: 1998-04-15
20 ; PRIOR APPLICATION NUMBER: US 06/04,583
21 ; PRIOR FILING DATE: 1997-04-16
22 ; NUMBER OF SEQ ID NOS: 2
23 ; SOFTWARE: PatentIn version 3.0
24 ; SEQ ID NO 2
25 ; LENGTH: 269
26 ; TYPE: PRT
27 ; ORGANISM: Homo sapiens
28 ; US-09-852-100A-2

```

```

Query Match: 100.0%; Score 202; DB 9; Length 269;
Best Local Similarity 100.0%; Pred. No. 26-190;
Matches 202; Conservative 0; Mismatches 0; Gaps 0;

CY 1 PPSGAPAEVATRLVGLVFWFSVTTGPMGAVATSAGGEESLKCCELVKQGVYICKDKPKIND 60
|||||
DB 68 PPSGAPAEVATRLVGLVFWFSVTTGPMGAVATSAGGEESLKCCELVKQGVYICKDKPKIND 127
|||||

CY 61 ATQPVNCTNTIAIVSCFPAPIITCKDSSGNETHFTGNVGFKEPISCRNNVNGYSYKVAV 120

```

```

Db 128 ATOEPVNCINYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFEPKPISCRNVNGYSYKAV 187
QY 121 ALSFLGWLGLADRYLGYPALGLLKFCTVPGCGISLIDFLISMQIVGSPDGSSTIYDY 180
Db 188 ALSFLGWLGLADRYLGYPALGLLKFCTVPGCGISLIDFLISMQIVGSPDGSSTIYDY 247
QY 181 YGTRLTRLSITNETFRKTOLYP 202
Db 248 YGTRLTRLSITNETFRKTOLYP 269

RESULT 2
US-09-833-503A-2
; Sequence 2, Application US/09833503A
; Patent No. US20020146760A1
; GENERAL INFORMATION:
; APPLICANT: Ozenberger, Bradley A
; APPLICANT: Kajkowski, Eileen M
; APPLICANT: Lo, Ching-Hsiung F
; APPLICANT: American Home Products Corporation
; TITLE OF INVENTION: No. US20020146760A1e1 G-Protein-Coupled Receptor-Like Proteins an
; TITLE OF INVENTION: Polynucleotides Encoded By Them, and Methods of Using
; FILE REFERENCE: AHP98165-00PC1
; CURRENT APPLICATION NUMBER: US/09/833,503A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/104,104
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 2
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-503A-2

Query Match 100.0%; Score 202; DB 10; Length 269;
Best Local Similarity 100.0%; Pred. No. 2e-190;
Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSGPSAPEAVTARLVGVLMFVSVITGPMGAVATSGAGSESLKCEDLKVGQYICKDKPKIND 60
Db 68 PSGPSAPEAVTARLVGVLMFVSVITGPMGAVATSGAGSESLKCEDLKVGQYICKDKPKIND 127
QY 61 ATOEPVNCINYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFEPKPISCRNVNGYSYKAV 120
Db 128 ATOEPVNCINYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFEPKPISCRNVNGYSYKAV 187
QY 121 ALSFLGWLGLADRYLGYPALGLLKFCTVPGCGISLIDFLISMQIVGSPDGSSTIYDY 180
Db 188 ALSFLGWLGLADRYLGYPALGLLKFCTVPGCGISLIDFLISMQIVGSPDGSSTIYDY 247
QY 181 YGTRLTRLSITNETFRKTOLYP 202
Db 248 YGTRLTRLSITNETFRKTOLYP 269

RESULT 3
US-10-199-881-2
; Sequence 2, Application US/10199881
; Publication No. US20030096356A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: No. US20030096356A1e1 G-Protein-Coupled Receptor-Like Proteins an
; TITLE OF INVENTION: Encoded by Them, and Methods of Using Same"
; FILE REFERENCE: AHP98165C1
; CURRENT APPLICATION NUMBER: US/10/199,881
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/US99/21621
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: US 90/833,5081
; PRIOR FILING DATE: 2001-12-04
```

```

; PRIOR APPLICATION NUMBER: US 60/104,104
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-881-2

Query Match 100.0%; Score 202; DB 15; Length 269;
Best Local Similarity 100.0%; Pred. No. 2e-190;
Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSGPSAPEAVTARLVGVLMFVSVITGPMGAVATSGAGSESLKCEDLKVGQYICKDKPKIND 60
Db 68 PSGPSAPEAVTARLVGVLMFVSVITGPMGAVATSGAGSESLKCEDLKVGQYICKDKPKIND 127
QY 61 ATOEPVNCINYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFEPKPISCRNVNGYSYKAV 120
Db 128 ATOEPVNCINYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFEPKPISCRNVNGYSYKAV 187
QY 121 ALSFLGWLGLADRYLGYPALGLLKFCTVPGCGISLIDFLISMQIVGSPDGSSTIYDY 180
Db 188 ALSFLGWLGLADRYLGYPALGLLKFCTVPGCGISLIDFLISMQIVGSPDGSSTIYDY 247
QY 181 YGTRLTRLSITNETFRKTOLYP 202
Db 248 YGTRLTRLSITNETFRKTOLYP 269

RESULT 4
US-10-226-956-288
; Sequence 288, Application US/10226956
; Publication No. US20030060399A1
; GENERAL INFORMATION:
; APPLICANT: Brophy, Colleen
; APPLICANT: Komalavilas, Padmini
; APPLICANT: Panitch, Alyssa
; APPLICANT: Joshi, Lokesh
; APPLICANT: Seal, Brandon L.
; TITLE OF INVENTION: REAGENTS AND METHODS FOR SMOOTH MUSCLE THERAPIES
; FILE REFERENCE: ASU-1061-US
; CURRENT APPLICATION NUMBER: US/10/226,956
; CURRENT FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: 60/314,535
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 320
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 288
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-226-956-288

Query Match 4.0%; Score 8; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 LFLGWLGA 131
Db 3 LFLGWLGA 10

RESULT 5
US-10-211-288-306
; Sequence 306, Application US/10211088
; Publication No. US20030104479A1
; GENERAL INFORMATION:
; APPLICANT: Bright, Gary R.
; APPLICANT: Premkumar, D. David
```

```
; APPLICANT: Chen, Yih-Tai
; TITLE OF INVENTION: NO. US20030104479A1el Fusion Proteins And Assays For Molecular Bi
; FILE REFERENCE: 01-1022-US
; CURRENT APPLICATION NUMBER: US/10/211,088
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/309,395
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/341,589
; PRIOR FILING DATE: 2001-12-13
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 306
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Protein-derived transport peptide
US-10-211-088-306

Query Match      4.0%; Score 8; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 LFLGWLGA 131
DB 3 LFLGWLGA 10

RESULT 6
US-09-915-914B-28
; Sequence 28, Application US/09015914B
; Publication No. US20030119725A1
; GENERAL INFORMATION:
; APPLICANT: Divita, Gilles
; APPLICANT: Fernandez, Joseph
; APPLICANT: Heitz, Frederic
; APPLICANT: Morris, May
; APPLICANT: Mery, Jean
; APPLICANT: Archdeacon, John
; APPLICANT: Horndorp, Kyle
; TITLE OF INVENTION: PEPTIDE-MEDIATED DELIVERY OF MOLECULES INTO CELLS
; FILE REFERENCE: AM-00105.P.1.1
; CURRENT APPLICATION NUMBER: US/09/9-5,914B
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: US 60/221,932
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic sequence
US-09-915-914B-28

Query Match      4.0%; Score 8; DB 11; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 LFLGWLGA 131
DB 3 LFLGWLGA 10

RESULT 7
US-09-785-802A-8
; Sequence 8, Application US/09785802A
; Patent No. US20020151004A1
; GENERAL INFORMATION:
; APPLICANT: Craig, Roger
; TITLE OF INVENTION: DELIVERY VEHICLES AND METHODS FOR USING THE SAME
; FILE REFERENCE: 11067/2035
```

```
; CURRENT APPLICATION NUMBER: US/09/785,802A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 09/748,06
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/748,789
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-785-802A-8

Query Match      4.0%; Score 8; DB 10; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 LFLGWLGA 131
DB 3 LFLGWLGA 10

RESULT 8
US-09-965-876A-2
; Sequence 2, Application US/09965876A
; Publication No. US20030096243A1
; GENERAL INFORMATION:
; APPLICANT: Busa, William B
; TITLE OF INVENTION: Methods and Reagents for Live-cell Gene Expression Quantifica
; FILE REFERENCE: 00-789-A
; CURRENT APPLICATION NUMBER: US/09/965,876A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/236,407
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 27
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-09-965-876A-2

Query Match      4.0%; Score 8; DB 11; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 LFLGWLGA 131
DB 3 LFLGWLGA 10

RESULT 9
US-09-915-914B-27
; Sequence 27, Application US/09915914B
; Publication No. US20030119725A1
; GENERAL INFORMATION:
; APPLICANT: Divita, Gilles
; APPLICANT: Fernandez, Joseph
; APPLICANT: Heitz, Frederic
; APPLICANT: Morris, May
; APPLICANT: Mery, Jean
; APPLICANT: Archdeacon, John
; APPLICANT: Horndorp, Kyle
; TITLE OF INVENTION: PEPTIDE-MEDIATED DELIVERY OF MOLECULES INTO CELLS
; FILE REFERENCE: AM-00105.P.1.1
; CURRENT APPLICATION NUMBER: US/09/915,914B
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: US 60/221,932
; PRIOR FILING DATE: 2000-07-31
```

; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic sequence
US-09-915-914B-27

Query Match 4.0%; Score 8; DB 11; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 LFLGWLGA 131
|||||
DB 3 LFLGWLGA 10

RESULT 10

US-10-077-555-8
; Sequence 8, Application US/10077555
; Publication No. US2003007289A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Rong-fu
; TITLE OF INVENTION: Use of Cell-Penetrating Peptides to Generate Antitumor Immunity
; FILE REFERENCE: P02373US/10200806
; CURRENT APPLICATION NUMBER: US/10/077,555
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/268,667
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-077-555-8

Query Match 4.0%; Score 8; DB 15; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 LFLGWLGA 131
|||||
DB 3 LFLGWLGA 10

RESULT 11

US-09-833-503A-6
; Sequence 6, Application US/09833503A
; Patent No. US20020146760A1
; GENERAL INFORMATION:
; APPLICANT: Ozenberger, Bradley A
; APPLICANT: Kajkowski, Eileen M
; APPLICANT: Lo, Ching-Hsiung F
; APPLICANT: American Home Products Corporation
; TITLE OF INVENTION: No. US20020146760A1el G-Protein-Coupled Receptor-Like Proteins and
; TITLE OF INVENTION: Polynucleotides Encoded By Them, and Methods of Using
; FILE REFERENCE: Same
; FILE REFERENCE: AHP98165-00PCT
; CURRENT APPLICATION NUMBER: US/09/833,503A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/104,104
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-833-503A-6

Query Match 4.0%; Score 9; DB 10; Length 221;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADREYLG 137
|||||
DB 174 GADREYLG 181

RESULT 12

US-09-992-600A-82
; Sequence 82, Application US/09992600A
; Publication No. US20030027161A1
; GENERAL INFORMATION:
; APPLICANT: Benjanin, Stephanie
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US4.DIV
; CURRENT APPLICATION NUMBER: US/09/992,600A
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: JPatent
; SEQ ID NO 82
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: 1..32
US-09-992-600A-82

Query Match 4.0%; Score 8; DB 11; Length 221;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADREYLG 137
|||||
DB 174 GADREYLG 181

RESULT 13

US-09-924-340-82
; Sequence 82, Application US/09924340
; Publication No. US20030027248A1
; GENERAL INFORMATION:
; APPLICANT: Benjanin, Stephanie
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US2.REG
; CURRENT APPLICATION NUMBER: US/09/924,340
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25

NUMBER OF SEQ ID NOS: 112
SOFTWARE: JPatent
SEQ ID NO 82
LENGTH: 221
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: 1..32
US-09-924-340-82

Query Match 4.0% Score 8; DB 11; Length 221;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADREYLG 137
DB 174 GADREYLG 181
|||||

RESULT 14

US-09-974-879-230
Sequence 230, Application US/09574879
Publication No. US2003002803A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 125 Human Secreted Proteins
FILE REFERENCE: P2029P2
CURRENT APPLICATION NUMBER: US/09/974,879
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: US 60/233,893
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: US 09/818,683
PRIOR FILING DATE: 2001-03-28
PRIOR APPLICATION NUMBER: US 09/305,736
PRIOR FILING DATE: 1999-05-05
PRIOR APPLICATION NUMBER: PCT/US98/23435
PRIOR FILING DATE: 1998-11-04
PRIOR APPLICATION NUMBER: US 60/064,911
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: US 60/064,912
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: US 60/064,983
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: US 60/064,900
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: US 60/064,988
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: US 60/064,987
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: US 60/064,908
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: US 60/064,984
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: US 60/064,985
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: US 60/066,094
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: US 60/066,100
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: US 60/065,089
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: US 60/066,095
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: US 60/066,090
PRIOR FILING DATE: 1997-11-17
NUMBER OF SEQ ID NOS: 611
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 230
LENGTH: 221
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:

NAME/KEY: SITE
LOCATION: (184)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-974-879-230

Query Match 4.0% Score 8; DB 11; Length 221;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADREYLG 137
DB 174 GADREYLG 181
|||||

RESULT 15

US-09-992-095B-82
Sequence 82, Application US/09992095B
Publication No. US20030157485A1
GENERAL INFORMATION:
APPLICANT: Benjamin, Stephanie
TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
FILE REFERENCE: 91,US5, DIV
CURRENT APPLICATION NUMBER: US/09/992,095B
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: US 60/924,340
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: PCT/IB01/01715
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: US 60/305,456
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: US 60/302,277
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/298,698
PRIOR FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: US 60/293,574
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 112
SOFTWARE: JPatent
SEQ ID NO 82
LENGTH: 221
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: 1..32
US-09-992-095B-82

Query Match 4.0% Score 8; DB 12; Length 221;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADREYLG 137
DB 174 GADREYLG 181
|||||

Search completed: September 26, 2003, 17:55:12
Job time : 71.4851 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:29:58 ; Search time 44.6979 Seconds
(without alignments)
434.602 Million cell updates/sec

Title: US-09-852-100A-2_COPY_68_269

Perfect score: 202
Sequence: 1 PSGPSAPEAVTARLVGVLMF.....TRLTSLSTNTEFRKIQLYP 202

Scoring table: OLLGO
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 95168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR.76.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	8	4.0	358	2 S23383	protein kinase (EC 2.7.1.37)
2	8	4.0	378	2 S22745	serine/threonine kinase
3	8	4.0	548	2 T16682	hypothetical prote
4	8	4.0	753	2 T28797	hypothetical prote
5	7	3.5	97	2 B97574	hypothetical prote
6	7	3.5	102	2 H70898	probable 2F protei
7	7	3.5	141	2 D75155	acetyl transferase
8	7	3.5	155	2 T48956	hypothetical prote
9	7	3.5	173	2 G71916	inorganic pyrophos
10	7	3.5	173	2 D64597	inorganic pyrophos
11	7	3.5	185	2 T20157	hypothetical prote
12	7	3.5	198	2 AB1598	conserved hypotet
13	7	3.5	198	2 AD1235	B. subtilis YncS p
14	7	3.5	202	2 A89911	conserved hypotet
15	7	3.5	204	2 F81445	type IV pilus asse
16	7	3.5	213	2 B86747	conserved hypotet
17	7	3.5	213	2 C95098	conserved hypotet
18	7	3.5	213	2 C97966	conserved hypotet
19	7	3.5	215	2 C69937	conserved hypotet
20	7	3.5	219	2 B75533	hypothetical prote
21	7	3.5	248	2 F71538	probable oxoacyl (
22	7	3.5	273	2 C71369	probable sugar ABC
23	7	3.5	293	2 D70136	conserved hypotet
24	7	3.5	337	2 T22128	hypothetical prote
25	7	3.5	317	2 A71237	hypothetical prote
26	7	3.5	343	2 B72507	hypothetical prote
27	7	3.5	349	2 B81736	probable 3'(2'),5'
28	7	3.5	384	2 AB0147	probable ABC trans
29	7	3.5	389	2 G61062	trans-sulfuratio

30 7 3.5 415 2 C83544 probable MFS trans
31 7 3.5 418 2 D86938 probable para-amin
32 7 3.5 419 1 J00346 ubiquinol-cytochrc
33 7 3.5 442 2 B75634 hypothetical prote
34 7 3.5 451 1 J00240 LIM kinase (EC 2.7
35 7 3.5 458 2 H70602 probable pabB prot
36 7 3.5 466 2 T44650 capsular polysacch
37 7 3.5 492 2 G75389 NADH2 dehydrogenas
38 7 3.5 495 2 A95984 probable xanthine
39 7 3.5 506 2 T12819 hypothetical prote
40 7 3.5 510 2 B71017 hypothetical prote
41 7 3.5 531 1 IKECB8 colicin B - Escher
42 7 3.5 540 2 C84744 probable psp19-lik
43 7 3.5 547 2 H75632 Na(+)-linked D-ala
44 7 3.5 575 2 T43400 myo-inositol trans
45 7 3.5 583 2 T25690 hypothetical prote

ALIGNMENTS

RESULT 1

S23383
protein kinase (EC 2.7.1.37) cdc2-related KKIALRE - human
C:Species: Homo sapiens (man)
C:Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 18-Jun-1999
C:Accession: S23383; S22744
C:Reviewer: M.; Enders, G.H.; Wu, C.L.; Su, L.K.; Gorka, C.; Nelson, C.; Harlow, E.
EMBL J. 11, 2909-2917, 1992
A:Title: A family of human cdc2-related protein kinases.
A:Reference number: S23382; MUID:92347325; PMID:1639063
A:Accession: S23383
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-358 <MEY>
A:Cross-references: EMBL:X66358; NID:q36614; PIDN:CAA47002.1; PID:q36615
C:Superfamily: Kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
E:3-278/Domain: protein kinase homology <KIN>
F:1-19/Region: protein kinase ATP-binding motif
F:54,52,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match: 4.0%; Score 8; DB 2; Length 358;
Best local similarity 100.0%; Pred. No. 3.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 138 YPALGLLK 145
DB 259 YPALGLLK 266

RESULT 2

S22745
serine/threonine protein kinase KKIALRE (EC 2.7.1.-) - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 19-Dec-1997
C:Accession: S22745
C:Reviewer: M.L.
submitted to the EMBL Data Library, May 1992
A:Reference number: S22743
A:Accession: S22745
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-376 <MEY>
A:Cross-references: EMBL:X66359
C:Genetics: 152/3; 170/3
C:Superfamily: Kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F:3-296/Domain: protein kinase homology <KIN>
F:11-19/Region: protein kinase ATP-binding motif

Query Match 4.0%; Score 8; DB 2; Length 376;

```

Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 YPALGLLK 145
    1111111
Db 277 YPALGLLK 284

RESULT 3
T16082
hypothetical protein F16H11.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000
C:Accession: T16082
R:Wu, X.
submitted to the EMBL Data Library, April 1996
A:Description: The sequence of C. elegans cosmid F16H11.
A:Reference number: Z18458
A:Accession: T16082
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-548 <WUX>
A:Cross-references: EMBL:U55376; NID:q1280130; PID:q1280130; PIDN:AAA98005.1; GSPDB:GN00169
A:Experimental source: strain Bristol N2; clone F16H11
C:Genetics:
A:Gene: CESP:F16H11.5
A:Map position: X
A:Introns: 71/1: 93/1; 187/3: 281/3; 316/3: 420/2; 460/1; 523/3; 545/3

Query Match 4.0%; Score 8; DB 2: Length 548;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 KOSSGNET 93
    1111111
Db 109 KOSSGNET 116

RESULT 4
T28787
hypothetical protein C41D11.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T28787
R:Gattung, S.; Maggi, L.
submitted to the EMBL Data Library, May 1997
A:Description: The sequence of C. elegans cosmid C41D11.
A:Reference number: Z20522
A:Accession: T28787
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-753 <GAT>
A:Cross-references: EMBL:AF003740; PIDN:AAC48141.1; GSPDB:GN00019; CESP:C41D11.5
A:Experimental source: strain Bristol N2; clone C41D11
C:Genetics:
A:Gene: CESP:C41D11.5
A:Map position: 1
A:Introns: 53/2: 81/3; 117/1: 256/3; 274/2: 357/3; 443/2; 485/3; 544/3; 585/3; 637/2

Query Match 4.0%; Score 8; DB 2: Length 753;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADHFYLG 137
    1111111
Db 398 GADHFYLG 405

RESULT 5
B97574
hypothetical protein AGR_C_3268 [imported] - Agrobacterium tumefaciens (strain C58, Cure
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002

```

```

C:Accession: B97574
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Oorollo, B.; Goic
A.: Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Iappas, C.; Markelz,
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: B97574
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-97 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK87547.1; PID:q15156883; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_3268
A:Map position: circular chromosome

Query Match 3.5%; Score 7; DB 2: Length 97;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 EAVTARL 14
    1111111
Db 47 EAVTARL 53

RESULT 6
H70898
Probable PE protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: H70898
R:Gale, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gorde
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Koltoyč
Raundread, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Salston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete gen
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: H70898
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-102 <COL>
A:Cross-references: GB:280108; GB:AL123456; NID:q3256012; PIDN:CAB02191.1; PID:e2655
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: PE

Query Match 3.5%; Score 7; DB 2: Length 102;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 EAVTARL 14
    1111111
Db 18 EAVTARL 24

RESULT 7
D75155
acetyl transferase PAB2081 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: D75155
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosom
A:Reference number: A75001
A:Accession: D75155
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-141 <KAW>
A:Cross-references: GB:AJ248284; GB:AL096836; NID:q5457730; PIDN:CAB49323.1; PID:e15
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB2081

```



```

Query Match          3.5%: Score 7; DB 2; Length 141;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Gaps 0;

Qy 149 VGFCGIG 155
Db 42 VGFCGIG 48

RESULT 8
T48956
hypothetical protein T15B3.160 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 13-Aug-2000
C:Accession: T48956
R:Jordan, N.; Bangert, S.; Wiedemann, R.; Voss, H.; Unselid, M.; Mewes, H.W.; Radd, S.;
submitted to the Protein Sequence Database, April 2000
A:Reference number: 225009
A:Accession: T48956
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-159 <J0R>
A:Cross-references: EMBL:AL163975; GSPDB:GN00061; ATSP:T15B3.160
A:Experimental source: cultivar Columbia; BAC clone T15B3
C:Genetics:
A:Gene: ATSP:T15B3.160
A:Map position: 3
A:Introns: 102/3; 115/1; 143/1
C:Superfamily: Arabidopsis thaliana hypothetical protein T15B3.160

Query Match          3.5%: Score 7; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 ALSIFLG 127
Db 94 ALSIFLG 100

RESULT 9
G71916
inorganic pyrophosphatase - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 22-Jun-1999
C:Accession: G71916
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, H.L.; Brown, E.D.; Dohg, P.C.; Smith, P.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, S.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: G71916
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-173 <ARN>
A:Cross-references: GB:AE001489; GB:AE001439; NID:94155102; PIDN:AA06646.1; PID:941551
A:Experimental source: strain J99
C:Genetics:
A:Gene: ppa
C:Superfamily: inorganic pyrophosphatase

Query Match          3.5%: Score 7; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 ARLVGL 18
Db 86 ARLVGL 92

RESULT 10
D64597

```

```

inorganic pyrophosphatase - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 22-Jun-1999
C:Accession: D64597
R:Tomb, J.F.; White, O.; Keriavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann,
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, E.; McI
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: D64597
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-173 <TOM>
A:Cross-references: GB:AE000576; GB:AE000511; NID:92313736; PIDN:AA07684.1; IID:92
C:Superfamily: inorganic pyrophosphatase

```

```

Query Match          3.5%: Score 7; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 ARLVGL 18
Db 86 ARLVGL 92

```

```

RESULT 11
T2057
hypothetical protein F58H1.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 23-Sep-2002
C:Accession: T20157; T22958
R:Holt, R.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z19230
A:Accession: T20157
A:Status: preliminary; translated from GB/EMBL/DD83
A:Molecule type: DNA
A:Residues: 1-185 <WIL>
A:Cross-references: EMBL:Z75012; PIDN:CAB01419.1; GSPDB:GN00023; CESP:F58H1.7
A:Experimental source: clone C52E4
R:Harris, B.
submitted to the EMBL Data Library, July 1996
A:Reference number: Z19642
A:Accession: T22958
A:Status: preliminary; translated from GB/EMBL/DD83
A:Molecule type: DNA
A:Residues: 1-185 <W12>
A:Cross-references: EMBL:Z75954; PIDN:CAB00110.1; GSPDB:GN00023; CESP:F58H1.7
A:Experimental source: clone F58H1
C:Genetics:
A:Gene: CFSP:F58H1.7
A:Map position: 5
A:Introns: 33/1; 77/1; 103/3; 128/1
C:Superfamily: LR11 protein; laminin-type EGF-like homology; LDL receptor ligand-bli
F:37-74/Domain: LDL receptor ligand-binding repeat homology <LDI>

```

```

Query Match          3.5%: Score 7; DB 2; Length 185;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 117 KVAVALS 123
Db 121 KVAVALS 127

```

```

RESULT 12
AB1598
Conserved hypothetical protein, B. subtilis YneS protein homolog lin123 [imported]
C:Species: Listeria innocua
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AB1598

```

A-Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A-Reference number: A89758; MUID:21311952; PMID:11418146
A-Accession: A89911

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:28:28 ; Search time 24.0681 Seconds
(without alignments);
394.688 Million cell updates/sec

Title: US-09-852-100A-2_COPY_68_269

Perfect score: 202
Sequence: 1 PSGPSAPENVATRLGVWF.....TRCTRLSITNETPKTQLYP 202

Scoring table: OLIGO

Gapop 60.0 ; Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	4.0	358	1 KRIA_HUMAN	Q00532 homo sapien
2	7	3.5	102	1 YP86_MYCTU	P71656 mycobacteri
3	7	3.5	168	1 YPAE_STRGC	Q9X972 streptococc
4	7	3.5	173	1 IPYR_HELPJ	Q9Z115 helicobacte
5	7	3.5	173	1 IPYR_HELPJ	P56153 helicobacte
6	7	3.5	198	1 YC84_LISKO	Q8Y713 listeria mo
7	7	3.5	198	1 YD23_LISIN	Q9ZC68 listeria in
8	7	3.5	202	1 YA35_STAEP	P59253 staphylococ
9	7	3.5	202	1 YD53_STAAM	Q99AC5 staphylococ
10	7	3.5	212	1 YC11_STRMU	P59255 streptococc
11	7	3.5	212	1 YC30_STR3	P59254 streptococc
12	7	3.5	213	1 Y851_SIRUN	Q54916 streptococc
13	7	3.5	213	1 Y908_STRPY	Q9A070 streptococc
14	7	3.5	213	1 YKAC_LACJA	Q9QW44 lactobacillus
15	7	3.5	215	1 YPUC_BACSU	P42978 bacillus su
16	7	3.5	224	1 CLSH_HUMAN	P56750 homo sapien
17	7	3.5	247	1 FABG_CHLTR	P38004 chlamydia t
18	7	3.5	292	1 SDB2_MOUSE	Q99J20 m.sc.musculi
19	7	3.5	307	1 RB33_CAEEL	Q20365 caenorhabdi
20	7	3.5	322	1 SC13_HUMAN	P55735 homo sapien
21	7	3.5	419	1 CYB_RHOVI	P81378 rhodopsin
22	7	3.5	469	1 CG51_HUMAN	Q9Y512 homo sapien
23	7	3.5	510	1 CEAR_ECOLI	P05819 escherichia
24	7	3.5	575	1 ITR1_SCHPO	Q10286 schizosacch
25	7	3.5	638	1 LIX2_RAT	P33670 rattus norv
26	7	3.5	663	1 CNQ2_BOVIN	Q33041 bos taurus
27	7	3.5	1039	1 AG43_ECOLI	P39180 escherichia
28	6	3.0	39	1 PSAG_PEA	P20120 pisum sativ
29	6	3.0	58	1 TY13_HUMAN	Q9BZ97 homo sapien
30	6	3.0	62	1 PUFK_RHOSH	P13402 rhodobacter
31	6	3.0	94	1 PTKB_ECOL6	Q8X7H5 escherichia
32	6	3.0	94	1 PTKB_ECOLI	P37188 escherichia
33	6	3.0	111	1 YG12_BACHD	Q9KCF9 bacillus ba

RESULT 1

ID	KXIA_HUMAN	STANDARD:	PRT:	358 AA.
AC	Q00532;			
DI	01-APR-1993 (Rel. 25, Created)			
DI	01-APR-1993 (Rel. 25, Last sequence update)			
DI	15-OCT-2001 (Rel. 40, Last annotation update)			
DE	Serine/threonine-protein kinase KXIALRE (EC 2.7.1.-) (Cyclin-dependent kinase-like 1).			
DE	kinase-like 1).			
GN	CDK1...			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
SK	11			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=92347325; PubMed=1639063;			
RA	Myerson M., Enders G.H., Wu C.-L., Su L.-K., Gorka C., Nelson C.,			
RA	Harlow E., Tsai L.-H.			
RT	"A family of human cdc2-related protein kinases.";			
RI	EMBL J. 11:2909-2917(1992).			
CC	-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.			
CC	CDK2/CDXX SUBFAMILY.			
CC	-----			
CC	This SW-SS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement. (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: X66358; CAA47002.1;			
DR	EMBL: X66359; CAA47002.1; JOINED.			
DR	PIR: S23383; S23383.			
DR	HSSP: P24941; LHCL.			
DR	Genew: HGNC:1781; CDK1.			
DR	SK: Q00532;			
DR	MIN: 603441;			
DR	GO: GO:0004693; F:Cyclin-dependent protein kinase activity; TAS.			
DR	GO: GO:0006468; P:protein amino acid phosphorylation; TAS.			
DR	GO: GO:0000074; P:regulation of cell cycle; TAS.			
DR	InterPro: IPR000719; Prot_kinase.			
DR	InterPro: IPR002290; Ser_thr_kinase.			
DR	PFam: PF00069; pkinase; 1.			
DR	ProDom: PD000001; Prot_kinase; 1.			
DR	SMART: SM00220; S_TKc; 1.			
DR	PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.			
DR	PROSITE: PS00108; PROTEIN_KINASE_ST; 1.			
DR	PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.			
DR	Transferase: Serine/threonine-protein kinase; ATP-binding.			
DR	DOMAIN 5 288			
DR	PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.			
DR	PROSITE: PS00108; PROTEIN_KINASE_ST; 1.			
DR	PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.			
DR	Transferase: Serine/threonine-protein kinase; ATP-binding.			
DR	NP_BIND 11 19			
DR	ATP (BY SIMILARITY).			
DR	BINDING 34 34			
DR	ATP (BY SIMILARITY).			
DR	ACT_SITE 127 127			
DR	BY SIMILARITY.			
DR	SEQUENCE 358 AA; 41834 MW; 88344321F24B77C6 CRC64;			

Q8X1P2 clostridium
P06017 trypanosoma
P06886 macropus ru
O43653 homo sapien
Q9X7A1 mycobacteri
O06327 mycobacteri
P81649 sus scrofa
Q8W226 methanopyru
Q8XAA3 escherichia
Q8IDA0 escherichia
P45465 escherichia
Q58625 methanococc

ALIGNMENTS

RESULT 1

KXIA_HUMAN
ID KXIA_HUMAN STANDARD: PRT: 358 AA.

AC Q00532;

DI 01-APR-1993 (Rel. 25, Created)

DI 01-APR-1993 (Rel. 25, Last sequence update)

DI 15-OCT-2001 (Rel. 40, Last annotation update)

DE Serine/threonine-protein kinase KXIALRE (EC 2.7.1.-) (Cyclin-dependent

kinase-like 1).

DE kinase-like 1).

GN CDK1...

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

SK 11

RP SEQUENCE FROM N.A.

RA MEDLINE=92347325; PubMed=1639063;

RA Myerson M., Enders G.H., Wu C.-L., Su L.-K., Gorka C., Nelson C.,

RA Harlow E., Tsai L.-H.

RT "A family of human cdc2-related protein kinases.";

RI EMBL J. 11:2909-2917(1992).

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC CDK2/CDXX SUBFAMILY.

This SW-SS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

EMBL: X66358; CAA47002.1;

EMBL: X66359; CAA47002.1; JOINED.

PIR: S23383; S23383.

HSSP: P24941; LHCL.

Genew: HGNC:1781; CDK1.

SK: Q00532;

MIN: 603441;

GO: GO:0004693; F:Cyclin-dependent protein kinase activity; TAS.

GO: GO:0006468; P:protein amino acid phosphorylation; TAS.

GO: GO:0000074; P:regulation of cell cycle; TAS.

InterPro: IPR000719; Prot_kinase.

InterPro: IPR002290; Ser_thr_kinase.

PFam: PF00069; pkinase; 1.

ProDom: PD000001; Prot_kinase; 1.

SMART: SM00220; S_TKc; 1.

PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.

PROSITE: PS00108; PROTEIN_KINASE_ST; 1.

PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.

Transferase: Serine/threonine-protein kinase; ATP-binding.

DOMAIN 5 288

PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.

PROSITE: PS00108; PROTEIN_KINASE_ST; 1.

PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.

Transferase: Serine/threonine-protein kinase; ATP-binding.

NP_BIND 11 19

ATP (BY SIMILARITY).

BINDING 34 34

ATP (BY SIMILARITY).

ACT_SITE 127 127

BY SIMILARITY.

SEQUENCE 358 AA; 41834 MW; 88344321F24B77C6 CRC64;

Query Match 4.0%; Score 8; DB 1; Length 358;
 Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 YPAUGLLK 245
 DE 259 YPAUGLLK 266

RESULT 2
 YDB6_MYCTU STANDARD; PRT: 102 AA.
 ID YDB6_MYCTU
 AC P71856;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein Rv1386 precursor.
 GN Rv1386 OR M1430 OR MTCY2184.03.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 CX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris R.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies K., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Horsley T., Jagels K., Krohn A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrett R.G.,
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 BAVTARL 14
 DE 18 BAVTARL 24

RESULT 3
 YPAE_STRUC STANDARD; PRT: 168 AA.
 ID YPAE_STRUC
 AC Q9X972;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein in parE 3' region (ORF2) (Fragment).
 OS Streptococcus gordonii Challis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 CX NCBI_TaxID=29390;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CHI:
 RA Vriesema A.J., Dankert J., Zaai S.A.;
 RT "Isolation and characterization of promoter regions from Streptococcus
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

Query Match 3.5%; Score 7; DB 1; Length 168;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
 DE 106 AVATSAG 112

RESULT 4
 IPYR_HELPJ STANDARD; PRT: 173 AA.
 ID IPYR_HELPJ
 AC Q92LL5;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-
 DE hydrolase) (PPase).
 GN PPA OR JHP0564.
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 OC Helicobacteraceae; Helicobacter.
 CX NCBI_TaxID=85963;
 RN [1]

```

RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9233882;
RA Alm R.A., Ling E.-S.H., Moir D.I., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang O., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT *Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.;
RL Nature 397:176-180(1999).
CC -!- CATALYTIC ACTIVITY: Diphosphate + H(2)O -> 2 phosphate.
CC -!- COFACTOR: REQUIRES THE PRESENCE OF DIVALENT METAL CATION.
CC MAGNESIUM CONFERS THE HIGHEST ACTIVITY. BINDS 4 DIVALENT CATIONS
CC PER SUBUNIT (BY SIMILARITY).
CC -!- SUBUNIT: Homohexamer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the Ppase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF001489; AAD06146.1; -
DR PIR: G71916; G71916.
DR HSP: P17288; 1FAJ.
DR HAMAP: MF_00209; -; 1.
DR InterPro: IPR001596; Pyrophosphatase.
DR Pfam: PF00719; Pyrophosphatase; 1.
DR ProDom: PD002014; Inorg_pphsp; 1.
DR PROSITE: PS00387; PPASE; 1.
KW Hydrolyase; Magnesium; Complete proteome.
FT ACT_SITE 28 28 BY SIMILARITY.
SQ SEQUENCE 173 AA; 19253 MW; 87B9B215E6FEHFC8 CMC64;

Query Match 3.5%; Score 7; DB 1; Length 173;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 ARLGVGL 18
DB 86 ARLGVGL 92

RESULT 5
IPYR_HELPY
ID IPYR_HELPY STANDARD; PRT; 173 AA.
AC P56153;
DT 01-NOV-1997 (Rel. 35; Created)
DT 01-NOV-1997 (Rel. 35; Last sequence update)
DI 28-FEB-2003 (Rel. 41; Last annotation update)
DE Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-
DE hydrolase) (PPase).
GN PPA OR HP0620.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kervilange A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Uterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Melman J.M., Fujii C., Bowman C., Watthey E., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,

```

```

RA Vector J.C.;
RT *The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.;
RI Nature 398:539-547(1997).
CC -!- CATALYTIC ACTIVITY: Diphosphate + H(2)O -> 2 phosphate.
CC -!- COFACTOR: REQUIRES THE PRESENCE OF DIVALENT METAL CATION.
CC MAGNESIUM CONFERS THE HIGHEST ACTIVITY. BINDS 4 DIVALENT CATIONS
CC PER SUBUNIT (BY SIMILARITY).
CC -!- SUBUNIT: Homohexamer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the Ppase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF000576; AAD07684.1; -
DR PIR: D64597; D64597.
DR HSP: P17288; 1FAJ.
DR TIGR: HP0620; -
DR HAMAP: MF_00209; -; 1.
DR InterPro: IPR001596; Pyrophosphatase.
DR Pfam: PF00719; Pyrophosphatase; 1.
DR ProDom: PD002014; Inorg_pphsp; 1.
DR PROSITE: PS00387; PPASE; 1.
KW Hydrolyase; Magnesium; Complete proteome.
FT ACT_SITE 28 28 BY SIMILARITY.
SQ SEQUENCE 173 AA; 19272 MW; 23A51C665A6E7F2 CRC64;

Query Match 3.5%; Score 7; DB 1; Length 173;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 ARLGVGL 18
DB 86 ARLGVGL 92

RESULT 6
YC84_L-SMO
ID YC84_L-SMO STANDARD; PRT; 198 AA.
AC O8Y713; O8KYA9;
DT 28-FEB-2003 (Rel. 41; Created)
DI 28-FEB-2003 (Rel. 41; Last sequence update)
DI 28-FEB-2003 (Rel. 41; Last annotation update)
DE Hypothetical protein Lmo1284.
GN Lmo1284.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=EGD;
RX MEDLINE=22035523; PubMed=12039883;
RA Lampidis R., Kostrewa D., Hof H.;
R: *Molecular characterization of the genes encoding DNA gyrase and
R: topoisomerase IV of Listeria monocytogenes.;
RL J. Antimicrob. Chemother. 49:917-924(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Deboux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Carrido P.,
RA Gautier L., Gobel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkak G.,

```

```

RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstrek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RT "Comparative genomics of Listeria species."
RL Science 294:849-852(2001).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AF084044; AAM48493.1; .
DR EMBL: AL591978; CAC99362.1; .
DR PIR: AD1235; AD1235; .
DR Listlist: LM01284; .
DR HAMAP: MF_01043; .; 1.
DR Pfam: PF02660; DUF205; 1.
DR TIGRFAMs: TIGR00023; TIGR00023; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 5 27 POTENTIAL.
FT TRANSMEM 85 107 POTENTIAL.
FT TRANSMEM 114 136 POTENTIAL.
FT TRANSMEM 156 178 POTENTIAL.
FT CONFLICT 27 29 IFY -> FST (IN REF. 1).
FT CONFLICT 45 45 N -> I (IN REF. 1).
FT CONFLICT 62 62 D -> V (IN REF. 1).
SQ SEQUENCE 198 AA; 21599 MW; 7807B5406DFC5CD1 CRC64;

Query Match 3.5%; Score 7; DB 1; Length 198;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
DB 109 AVATSAG 115

RESULT 7
YD23 LISIN STANDARD; PRT; 198 AA.
AC Q92C58; 28-FEB-2003 (Rel. 41, Created)
DI 28-FEB-2003 (Rel. 41, Last sequence update)
DI 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein: Lin1323.
GN LIN1323.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Franche P., Bloeker H., Brandt P., Chakraborty T.,
RA Baquero F., Berche P., Boue E., de Bataillon A., Debois P.,
RA Charbit A., Chetoui F., Couve E., Duchaud E., Durant L., Dussurget O.,
RA Domann E., Dominguez-Bernal G., Duclaud E., Garrido P.,
RA Entian K.-D., Fshih H., Garcia-del Portillo F., Hain T., Haul C., Jackson D.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haul C., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapatk S.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstrek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RT "Comparative genomics of Listeria species."
RL Science 294:849-852(2001).

```

```

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AL596168; CAC96554.1; .
DR PIR: AB1598; AB1598; .
DR Listlist: LIN01323; .
DR HAMAP: MF_01043; .; 1.
DR InterPro: IPR003811; DUF205; 1.
DR Pfam: PF02660; DUF205; 1.
DR TIGRFAMs: TIGR00023; TIGR00023; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 5 27 POTENTIAL.
FT TRANSMEM 85 107 POTENTIAL.
FT TRANSMEM 114 136 POTENTIAL.
FT TRANSMEM 156 178 POTENTIAL.
SQ SEQUENCE 198 AA; 21632 MW; B161D1055R203406 CRC64;

Query Match 3.5%; Score 7; DB 1; Length 198;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
DB 109 AVATSAG 115

RESULT 8
YA35 STAEPP STANDARD; PRT; 202 AA.
AC P59253; 28-FEB-2003 (Rel. 41, Created)
DI 28-FEB-2003 (Rel. 41, Last sequence update)
DI 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein SE1035.
GN SE1035.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,
RA Chen Z., Wen Y.;
RA Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AE016747; AA004632.1; .
DR HAMAP: MF_01043; .; 1.
DR Pfam: PF02660; DUF205; 1.
DR TIGRFAMs: TIGR00023; TIGR00023; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 4 26 POTENTIAL.
FT TRANSMEM 52 74 POTENTIAL.
FT TRANSMEM 84 106 POTENTIAL.
FT TRANSMEM 118 140 POTENTIAL.
FT TRANSMEM 141 161 POTENTIAL.

```

FT TRANSMEM 162 182 POTENTIAL.
SQ SEQUENCE 202 AA; 22105 MW; 6516AFA75DD72F0 CRC64;

Query Match 3.5%; Score 7; DB 1; Length 202;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
11111111
DB 113 AVATSAG 119

RESULT 9

YD53_STAM
ID YD53_STAM STANDARD; PRT; 202 AA.
AC Q99UC5;
DT 28-FEB-2003 (Rel. 41, Created)
DI 28-FEB-2003 (Rel. 41, Last sequence update)
DI 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein SSV1353/SA1187/MW1240.
GN SSV1353 OR SA1187 OR MW1240.
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus (strain MW21).
OC Bacteria; Firmicutes; Bacillales; Staphylococcaceae.
OX NCBI_TaxID=158878, 158879, 196620;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=MW50 / ATCC 700699, and N315;
RC MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama Y., Baba T., Yuzawa H., Kobayashi T.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-O., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hoshoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Heyashi H., Hiramatsu K.,
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus".
RL Lancet 357:1225-1240(2001).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=MW2;
RC MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamanoto K., Hiramatsu K.,
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA".
RL Lancet 359:1819-1827(2002).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL: AP003362; BAB57515.1;
CC PIR: A89911; A89911.
CC HAMAP: MF_01043; -; 1.
CC InterPro: IPR003811; DUF205.
CC Pfam: PF02660; DUF205; 1.
CC TIGRFAMs: TIGR00023; TIGR00023; 1.
CC Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 4 26 POTENTIAL.
FT TRANSMEM 52 74 POTENTIAL.
FT TRANSMEM 84 106 POTENTIAL.

FT TRANSMEM 118 140 POTENTIAL.
FT TRANSMEM 141 161 POTENTIAL.
FT TRANSMEM 162 182 POTENTIAL.
SQ SEQUENCE 202 AA; 22232 MW; A9DA126B5731749C CRC64;

Query Match 3.5%; Score 7; DB 1; Length 202;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
11111111
DB 113 AVATSAG 119

RESULT 10

YC11_STRMU
ID YC11_STRMU STANDARD; PRT; 212 AA.
AC P59255;
DT 28-FEB-2003 (Rel. 41, Created)
DI 28-FEB-2003 (Rel. 41, Last sequence update)
DI 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein SMU.1211.
GN SMU.1211.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=UA159 / ATCC 700610 / Serotype C;
RC MEDLINE=22295063; PubMed=12397186;
RA Adic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.,
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen".
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL: AE014958; AAN58897.1;
CC HAMAP: MF_01043; -; 1.
CC Pfam: PF02660; DUF205; 1.
CC TIGRFAMs: TIGR00023; TIGR00023; 1.
CC Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 4 23 POTENTIAL.
FT TRANSMEM 69 91 POTENTIAL.
FT TRANSMEM 111 133 POTENTIAL.
FT TRANSMEM 140 162 POTENTIAL.
FT TRANSMEM 166 183 POTENTIAL.
SQ SEQUENCE 212 AA; 23417 MW; D4DA5F4270C5F5B CRC64;

Query Match 3.5%; Score 7; DB 1; Length 212;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
11111111
DB 105 AVATSAG 111

RESULT 11

YC30_STRAS
ID YC30_STRAS STANDARD; PRT; 212 AA.
AC P59254;

```

DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein gbs1230/SAG1155.
CN GBS1230 OR SAG1155.
OS Streptococcus agalactiae (serotype III), and
OC Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=216495, 216466;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=MEM316 / Serotype III;
RX MEDLINE=22242503; PubMed=12354221;
RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Frangoul L.,
RA Meadek T., Zouine M., Couve E., Talloui L., Poyart C., Troadec P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease.";
RL Mol. Microbiol. 45:1499-1513(2002).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=2222988; PubMed=12200547;
RA Tettelin H., Masignani V., Cieslewicz M.J., Fisen J.A., Peterson S.,
RA Wessels M.R., Paulsen I.T., Nelson K.E., Margalit I., Read T.D.,
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.A.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Willigan S.,
RA Carthy A.A., Cline R.T., Van Aken S.E., Gill J., Scariselli M., Mora M.,
RA Iacobini E.T., Brettoni C., Galili G., Mariani M., Veqni F., Malone D.,
RA Rinaldo D., Rappuoli R., Telford J.L., Kaspar D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE UPF0078 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL766849; CAD46889.1; -
DR EMBL: AE014244; AAN0037.1; -
DR SAG1230; gbs1230; -
DR HAMAP: MF_01043; -; 1.
DR Pfam: PF02660; DUF205; 1.
DR TIGRFAMs: TIGR00023; TIGR00023; 1.
DR TIGRFAMs: TIGR00023; Transmembrane; Complete proteome.
FT TRANSMEM 4 23 POTENTIAL.
FT TRANSMEM 69 91 POTENTIAL.
FT TRANSMEM 113 135 POTENTIAL.
FT TRANSMEM 142 164 POTENTIAL.
FT TRANSMEM 168 185 POTENTIAL.
SQ SEQUENCE 212 AA; 23403 MW; 2BFE4A06F04E1657 CRC64;

Query Match 3.5%; Score 7; Ds 1; Length 212;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative C; Mismatches C; Indels 0; Gaps C;

QY 30 AVATSAG 36
DB 105 AVATSAG 111
|||||||

RESULT 12
Y851_STRPN

```



```
FT TRANSMEM 2 22 POTENTIAL.
FT TRANSMEM 70 90 POTENTIAL.
FT TRANSMEM 112 132 POTENTIAL.
FT TRANSMEM 143 163 POTENTIAL.
FT TRANSMEM 164 184 POTENTIAL.
FT TRANSMEM 168 188 POTENTIAL.
FT CONFLICT 168 188 S -> F (IN REF. 1).
SQ SEQUENCE 213 AA: 22928 MW: 22CB089C17750818 CKC64;

Query Match 3.5% Score 7; DB 1: Length 213;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
DB 105 AVATSAG 111

RESULT 13
Y908_STRPY STANDARD; PRT; 213 AA.
ID Y908_STRPY STANDARD; PRT; 213 AA.
AC Q9A070; Q8K708;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein SPY0908/spyM3_0623/SPS1230/spyM8_0566.
GN SPY0908 OR SPYM3_0623 OR SPS1230 OR SPYM18_0566.
OS Streptococcus pyogenes,
OS Streptococcus pyogenes (serotype M3), and
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314, 198466, 186103;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN-SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White C.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.,
RA "Complete genome sequence of an M1 strain of Streptococcus pyogenes."
RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
RN [2];
RP SEQUENCE FROM N.A.
RC STRAIN-MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206;
RA Beres S.B., Sylva G., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins J.F.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence."
RT Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
RN [3];
RP SEQUENCE FROM N.A.
RC STRAIN-SST-1 / Serotype M3;
RX Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
RA Yamazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,
RA Hayashi H., Hamada S.;
RT "The genome of invasive Streptococcus pyogenes: a comparative analysis
RT of S. pyogenes SST-1, SF370 and MGAS232."
RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [4];
RP SEQUENCE FROM N.A.
RC STRAIN-MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins J.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
```

```

DR InterPro: IPR003811; DUF205.
DR Pfam: PF02660; DUF205; 1.
DR TIGRFAMs: TIGR00023; TIGR00023; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 2 24 POTENTIAL.
FT TRANSMEM 44 63 POTENTIAL.
FT TRANSMEM 70 92 POTENTIAL.
FT TRANSMEM 112 134 POTENTIAL.
FT TRANSMEM 141 163 POTENTIAL.
FT TRANSMEM 167 184 POTENTIAL.
SQ SEQUENCE 213 AA; 23362 MW; 967258F000F3AE3 CRC64;

Query Match 3.5%; Score 7; DB 1; Length 215;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 AVATSAG 36
DB 106 AVATSAG 112

RESULT 15
YPJC_BACSU STANDARD; PRT; 215 AA.
AC P42978;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DI 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein ypJC.
GN ypJC OR JOJC.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / Marburg;
RX MEDLINE=96349105; PubMed=8760912;
RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
RA Serror P.;
RT "Sequence analysis of the Bacillus subtilis chromosome region between:
RL the serA and kds loci cloned in a yeast artificial chromosome.";
RN Microbiology 142:2005-2016(1996).
RP [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.C., Bessieres P., Bolotin A., Borcher S.,
RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.J.,
RA Entian K.D., Ervington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi S.,
RA Guiseppi G., Guy B.J., Haga K., Haeck J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Potworlik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Selde B.,
RA Sorokin A., Tacconi E., Takagi I., Takahashi H., Takemaru K.,
RA Takeuchi M., Takashi A., Takata T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzengger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,

```

```

RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RL subtilis.";
RL Nature 390:249-256(1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: TO B.SUBTILIS YITT AND YQPI.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L38424; AAA92872.1; -
DR EMBL: L47709; AAB38440.1; -
DR EMBL: Z99115; CAB14167.1; -
DR PIR: C69937; C69937.
DR Subtilist: BGL1209; ypJC.
DR InterPro: IPR003740; DUF151.
DR Pfam: PF02588; DUF161; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 31 51 POTENTIAL.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 104 124 POTENTIAL.
SQ SEQUENCE 215 AA; 23582 MW; D314CF7225F8A983 CRC64;

Query Match 3.5%; Score 7; DB 1; Length 215;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 VALSLFL 126
DB 10 VALSLFL 16

Search completed: September 26, 2003, 17:37:19
Job time : 26.0681 secs

```

GenCore version: 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:29:43 : Search time 104.868 Seconds
(without alignments)
497.069 Million cell updates/sec

Title: US-09-852-100a-2_COPY_68_269

Perfect score: 202

Sequence: 1 PSGPSAPEAVTARLVGVLMF.....TLRLSLTNETFRKTKQLYP 202

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

SPREMBL_23:*

- 1: sp.archaea:*
- 2: sp.bacteria:*
- 3: sp.fungi:*
- 4: sp.human:*
- 5: sp.invertebrate:*
- 6: sp.mammal:*
- 7: sp.mhc:*
- 8: sp.organella:*
- 9: sp.phage:*
- 10: sp.plant:*
- 11: sp.podent:*
- 12: sp.virus:*
- 13: sp.vertebrate:*
- 14: sp.unclassified:*
- 15: sp.rvirus:*
- 16: sp.bacteriap:*
- 17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	202	100.0	207	4 Q9BX74	Q9BX74 homo sapien
2	99	49.0	208	1 Q99MB3	Q99MB3 mus musculu
3	8	4.0	80	4 Q9H046	Q9H046 homo sapien
4	8	4.0	178	5 Q9W2H1	Q9W2H1 drosophila
5	8	4.0	195	5 Q95025	Q95025 caenorhabdi
6	8	4.0	206	2 Q93164	Q93164 rhodovulum
7	8	4.0	221	4 Q9H651	Q9H651 homo sapien
8	8	4.0	230	11 Q90156	Q90156 mus musculu
9	8	4.0	247	4 Q9B3N9	Q9B3N9 homo sapien
10	8	4.0	261	11 Q9B383	Q9B383 mus musculu
11	8	4.0	284	5 Q9G4H5	Q9G4H5 drosophila
12	8	4.0	548	5 Q19496	Q19496 caenorhabdi
13	8	4.0	1638	5 Q9VCW7	Q9VCW7 drosophila
14	7	3.5	38	11 Q8CAM1	Q8CAM1 mus musculu
15	7	3.5	97	16 Q8U561	Q8U561 agrobacteri
16	7	3.5	132	2 Q9S3F2	Q9S3F2 helicobacte

17	7	3.5	132	2 Q8VLC3	Q8VLC3 helicobacte
18	7	3.5	132	2 Q8VNT8	Q8VNT8 helicobacte
19	7	3.5	132	2 Q9S3F1	Q9S3F1 helicobacte
20	7	3.5	132	2 Q8VLG4	Q8VLG4 helicobacte
21	7	3.5	132	2 Q9S3F4	Q9S3F4 helicobacte
22	7	3.5	132	2 Q8VNF9	Q8VNF9 helicobacte
23	7	3.5	132	2 Q9S3F0	Q9S3F0 helicobacte
24	7	3.5	132	2 Q9S3F3	Q9S3F3 helicobacte
25	7	3.5	132	2 Q8V5F8	Q8V5F8 helicobacte
26	7	3.5	132	2 Q9R3S8	Q9R3S8 helicobacte
27	7	3.5	132	2 Q9S3E9	Q9S3E9 helicobacte
28	7	3.5	141	2 Q9AC92	Q9AC92 mycobacteri
29	7	3.5	141	2 Q8G656	Q8G656 mycobacteri
30	7	3.5	141	17 Q9VIM6	Q9VIM6 pyrococcus
31	7	3.5	150	16 Q8EV17	Q8EV17 mycoplasma
32	7	3.5	159	10 Q9LXV9	Q9LXV9 arabidopsis
33	7	3.5	159	16 Q8F099	Q8F099 corynebacte
34	7	3.5	173	2 Q6VN51	Q6VN51 helicobacte
35	7	3.5	173	2 Q93CN1	Q93CN1 helicobacte
36	7	3.5	173	2 Q8VN49	Q8VN49 helicobacte
37	7	3.5	173	2 Q8GK71	Q8GK71 helicobacte
38	7	3.5	173	2 Q8GK70	Q8GK70 helicobacte
39	7	3.5	173	2 Q8GK68	Q8GK68 helicobacte
40	7	3.5	173	2 Q8GK67	Q8GK67 helicobacte
41	7	3.5	173	2 Q8GK66	Q8GK66 helicobacte
42	7	3.5	173	2 Q8GK65	Q8GK65 helicobacte
43	7	3.5	173	2 Q8GK64	Q8GK64 helicobacte
44	7	3.5	173	2 Q8GK63	Q8GK63 helicobacte
45	7	3.5	173	2 Q8GK62	Q8GK62 helicobacte

ALIGNMENTS

RESULT 1

Q9BX74 PRELIMINARY: PRT: 207 AA.
AC Q9BX74;
DT 01-JUN-2001 (TRENBLREL. 17, Created)
DT 01-JUN-2001 (TRENBLREL. 17, Last sequence update)
DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
DE Beta-amyloid binding protein.
GN BDP.
GS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
GX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278845;
RA Kajkowski E.M., Lo C.F., Ning X., Walker S., Sofia H.J., Wang W.,
RA Edris W., Chanda P., Wagner E., Vile S., Ryan K., McHenry-Rinde R.,
RA Smith S.C., Wood A., Rhodes K.J., Kennedy J.D., Bard J.,
RA Jacobsen J.S., Ozenberger B.A.;
FT "Beta-Amyloid Peptide-Induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module.";
RL J. Biol. Chem. 276:16748-16756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Testis;
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF353990; AAK35084.1; -;
DR EMBL; BC029486; AAH29486.1; -;
KW Signal.
FT SIGNAL. 1 37 POTENTIAL.
SQ SEQUENCE 207 AA: 22326 MW: A5590FD7AECDF292 CRC64;

Query Match 100.0%; Score 202; DB 4; Length 207;
Best Local Similarity 100.0%; Pred. No. 4.3e-205;
Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSGPSAPEAVTARLVGVLMFVSVTTGPWGAATVATSGAESLKCEDLVGVYICKDKPKIND 60


```

RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guerin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Y. C., Lewis S.E., Rubin G.M., Celisner S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AEO03453; AAF46720.1; -
DR EMBL: AY061343; AAL28891.1; -
DR FlyBase: FBgn034626; CG10795.
SQ SEQUENCE 178 AA; 19896 MW; 17C41166607ACC03 CRC64;

Query Match 4.0%; Score 8; DB 5; Length 178;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 DREYLGYP 139
DB 108 DREYLGYP 115
|||||||

RESULT 5
Q95Q25 PRELIMINARY; PRT; 195 AA.
AC Q95Q25;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 21.2 kDa protein.
GN C41D11.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorca; Rhabditida; Rhabditidae;
OC Rhabditidae; Peleoderinae; Caenorhabditis.
CX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998);
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gattung S., Magoi L.;
RT "The sequence of C. elegans cosmid C41D11."
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF003740; AAL03031.1; -
DR WormPeP: C41D11.9; CE29489.
KW Hypothetical protein.
SQ SEQUENCE 195 AA; 21203 MW; 35945F407F184DAE CRC64;

Query Match 4.0%; Score 8; DB 5; Length 195;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADRFYLG 137
DB 148 GADRFYLG 155
|||||||

RESULT 6
Q93I64 PRELIMINARY; PRT; 206 AA.
AC Q93I64;

```

```

DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE BCHO (Fragment).
GN BCHO.
OS Rhodovulum sulfidophilum (Rhodobacter sulfidophilus).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Rhodovulum.
CX NCBI_TaxID=35806;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W4;
RA Yoshida M., Masuda S., Nagashima K.V., Vermeiglio A., Shinada K.,
RA Matsuura K.;
RT "In vitro and in vivo electron transfer to the trHEME cytochrome
RT s-subunit bound to the photosynthetic reaction center complex in the
RT purple bacterium Rhodovulum sulfidophilum."
RI Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB050579; BAB70491.1; -
DR InterPro: IPR000073; A/B_hydrolase.
DR InterPro: IPR003089; AB_hydrolase.
DR InterPro: IPR000379; Ser_estrs_site.
DR Pfam: PF00561; abhydrolase; 1.
DR PRINTS: PR00111; ABHYDROLASE.
FI NON_TER 1
SQ SEQUENCE 206 AA; 21500 MW; 8F6713A03BB5B076 CRC64;

Query Match 4.0%; Score 8; DB 2; Length 206;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 APEAVTAR 13
DB 190 APEAVTAR 197
|||||||

RESULT 7
Q9H651 PRELIMINARY; PRT; 221 AA.
AC Q9H651;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ22604 (BBP-like protein 2).
GN BL22.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUP=Small intestine;
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Uta T.,
RA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isonaga T., Sugano S.;
RT "NEDO human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kajkowski E.M., Lo C.F., Ning X., Walker S., Sofia H.J., Wang W.,
RA Edris W., Chanda P., Waquer E., Vile S., Ryan K., McHenry-Rinde B.,
RA Smith S.C., Wood A., Rhodes K.J., Kennedy J.D., Bard J.,
RA Jacobsen J.S., Ozenberger B.A.;
RT "Beta-amyloid peptide-induced apoptosis regulated by a novel protein
RT containing a G protein activation module."
RL J. Biol. Chem. 0:0-0(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK026257; BAB15415.1; -
DR EMBL: AF353992; AAK35066.1; -

```

```

DR EMBL; BC008973; AAH08873.1; -.
KW Hypothetical protein.
SQ SEQUENCE 221 AA; 24410 MW; 92151D6E6F636D74 CRC64;

Query Match
Best Local Similarity 4.0%; Score 8; DB 4; Length 221;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADRFYLG 137
DB 174 GADRFYLG 181

RESULT 8
Q9D156 PRELIMINARY; PRT; 230 AA.
AC Q9D156;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DI 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE Similar to hypothetical protein FLJ22604.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TSSUE=Colon;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006150; AAH06150.1; -.
KW Hypothetical protein.
SQ SEQUENCE 247 AA; 27161 MW; CE1D0D9C53DDF73C CRC64;

Query Match
Best Local Similarity 4.0%; Score 8; DB 4; Length 247;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADRFYLG 137
DB 200 GADRFYLG 207

RESULT 10
CE5J83 PRELIMINARY; PRT; 261 AA.
AC CE5J83;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DI 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Similar to Bsp-like protein 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT *Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.
RL Nature 420:563-573(2002).
DR EMBL; AK077858; BAC37037.1; -.
SQ SEQUENCE 261 AA; 28880 MW; 70346780D3CF5CDB CRC64;

Query Match
Best Local Similarity 4.0%; Score 8; DB 11; Length 261;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADRFYLG 137
DB 214 GADRFYLG 22;

RESULT 11
Q9U4H5 PRELIMINARY; PRT; 294 AA.
AC Q9U4H5; Q9W361;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DI 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE Similar to hypothetical protein FLJ22604.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT *Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.
RL Nature 420:563-573(2002).
DR EMBL; AK077858; BAC37037.1; -.
SQ SEQUENCE 261 AA; 28880 MW; 70346780D3CF5CDB CRC64;

Query Match
Best Local Similarity 4.0%; Score 8; DB 11; Length 230;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADRFYLG 137
DB 183 GADRFYLG 190

RESULT 9
Q9BRN9 PRELIMINARY; PRT; 247 AA.
AC Q9BRN9;

```

```

DR EMBL; BC008973; AAH08873.1; -.
KW Hypothetical protein.
SQ SEQUENCE 221 AA; 24410 MW; 92151D6E6F636D74 CRC64;

Query Match
Best Local Similarity 4.0%; Score 8; DB 4; Length 221;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADRFYLG 137
DB 174 GADRFYLG 181

RESULT 8
Q9D156 PRELIMINARY; PRT; 230 AA.
AC Q9D156;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DI 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE Similar to hypothetical protein FLJ22604.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TSSUE=Colon;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006150; AAH06150.1; -.
KW Hypothetical protein.
SQ SEQUENCE 247 AA; 27161 MW; CE1D0D9C53DDF73C CRC64;

Query Match
Best Local Similarity 4.0%; Score 8; DB 4; Length 247;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADRFYLG 137
DB 200 GADRFYLG 207

RESULT 10
CE5J83 PRELIMINARY; PRT; 261 AA.
AC CE5J83;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DI 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Similar to Bsp-like protein 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT *Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.
RL Nature 420:563-573(2002).
DR EMBL; AK077858; BAC37037.1; -.
SQ SEQUENCE 261 AA; 28880 MW; 70346780D3CF5CDB CRC64;

Query Match
Best Local Similarity 4.0%; Score 8; DB 11; Length 261;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADRFYLG 137
DB 214 GADRFYLG 22;

RESULT 11
Q9U4H5 PRELIMINARY; PRT; 294 AA.
AC Q9U4H5; Q9W361;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DI 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE Similar to hypothetical protein FLJ22604.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT *Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.
RL Nature 420:563-573(2002).
DR EMBL; AK077858; BAC37037.1; -.
SQ SEQUENCE 261 AA; 28880 MW; 70346780D3CF5CDB CRC64;

Query Match
Best Local Similarity 4.0%; Score 8; DB 11; Length 230;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GADRFYLG 137
DB 183 GADRFYLG 190

RESULT 9
Q9BRN9 PRELIMINARY; PRT; 247 AA.
AC Q9BRN9;

```

```

DQ 237 GADRFYLG 244

RESULT 12
QI3496 PRELIMINARY; PRT; 548 AA.
ID QI3496
AC QI3496;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 63.5 kDa protein.
F16H11.5
GN Caenorhabditis elegans.
OS Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdiida; Rhabdioida;
OC Rhabditia; Pezodoriinae; Caenorhabditis.
OX NCBI_TaxID=6239;
(1)
KN SEQUENCE FROM N.A.
RP STRAIN-Bristol M2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN 1]
RC "The sequence of C. elegans cosmid F16H11.";
RA Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
(3]
KN SEQUENCE FROM N.A.
RP STRAIN-Bristol M2;
RX STRAIN-Bristol M2;
RA Wu X.;
RN 1]
RC "Direct Submission.";
RA Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RL SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC 1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
EMBL: U55376; AAA98005.1; ..
DR HSP: P20393; 1A6Y.
DR WormPep: F16H11.5; C504400.
DR InterPro: IPR000536; Hormone_rec_lig.
DR InterPro: IPR001628; Znfc4steroid.
DR Pfam: PF00104; hormone_rec; 1.
DR Pfam: PF00105; zf-C4; 1.
DR PRINTS: PR00047; STROIDFINGER.
DR ProDom: PD000035; Znfc4steroid; 1.
DR SMART: SM00430; HOLI; 1.
DR SMART: SM00399; Znfc4; 1.
KW Hypothetical protein; DNA-binding; Metal-binding; Nuclear protein;
Receptor; Transcription; Transcription regulation; Zinc; Zinc-finger
SQ SEQUENCE 548 AA; 63545 MW; 3B5C51C6850BB96 CRC64;

Query Match 4.0%; Score 6; DB 5; Length 546;
Rest local similarity 100.0%; Pred No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 66 KDSSGNET 93
II:IIII:
DB 109 KDSSGNET 116

RESULT 13
Q9VCW7 PRELIMINARY; PRT; 1638 AA.
ID Q9VCW7
AC Q9VCW7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE CG6954 protein.
GN Drosophila melanogaster (Fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

```

OC Neoptera: Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea: Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Koskins R.A., Gallie R.E.,
 RA Sutton G.G., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang C., Chen L.X.,
 RA Brandon R.C., Rogers J.H.C., Blaise R.G., Champ C., Pfaffner B.E.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Adaya A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Balieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan N.R., Bouck J., Brockstein P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Danke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann N.,
 RA Foster C., Gabrielson A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.C., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Itoqwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reider K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Stimpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.C.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003740; AAF56038.1; ..
 DR FlyBase: FBgn0039000; CG6954.
 DR InterPro: IPR002110; ANK.
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR000159; RA_domain.
 DR Pfam: PF00023; ank; 2.
 DR Pfam: PF00041; fn3; 1.
 DR Pfam: PF00788; RA; 1.
 DR SMART: SM00248; ANK; 1.
 DR SMART: SM00060; FN3; 1.
 DR SMART: SM00314; RA; 1.
 DR PROSITE: PS50088; ANK_REPEAT; 1.
 DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
 KW ANK repeat. Repeat.
 SQ SEQUENCE 1638 AA: 180548 MW: 3CC84F9A24B9F5DC CRC64;

Query Match 4.0%; Score 8; DB 5; Length 1638;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 AVATSAGG 37
 DB 1241 AVATSAGG 1248
 |||||

RESULT 14

Q8CAM1

PRELIMINARY:

PRT:

38 AA.

Q8CAM1;
 CT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE LIM motif-containing protein kinase 2 (fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=CS7BL/6J; TISSUE=Hypothalamus;
 RC MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium.
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RI Nature 420:563-573(2002).
 DR EMBL: AK038530; BAC30030.1; ..
 FT NON_TER 1
 SQ SEQUENCE 38 AA: 4188 MW: 2DB363A494415D42 CRC64;
 Query Match 3.5%; Score 7; DB 11; Length 38;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 121 ALSFLG 127
 DB 3 ALSFLG 9
 |||||

RESULT 15

Q8U581

PRELIMINARY:

PRT:

97 AA.

AC Q8U581
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE AGR_C_3268p.
 GN AGR_C_3268.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 CC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Grollio B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Houmlel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Dougherty D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58";
 RL Science 294:2323-2328(2001).
 DR EMBL: AE008099; AAK87547.1; ..
 SQ SEQUENCE 97 AA: 10042 MW: 6B2B2E03D4137F36 CRC64;

Query Match 3.5%; Score 7; DB 16; Length 97;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 EAVTARL 14
 DB 47 EAVTARL 53
 |||||

Search completed: September 26, 2003, 17:41:35
 Job time : 106.868 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2003, 17:22:39 : Search time 97 seconds
(without alignments)
715.630 Million cell updates/sec

Title: US-09-852-100A-2

Perfect score: 1439

Sequence: 1 MHLLKGNPVPRAHQKNT.....TRLTKLSITNEFRKIQ:YP 269

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp.archaea:*
- 2: sp.bacteria:*
- 3: sp.fungi:*
- 4: sp.human:*
- 5: sp.invertebrate:*
- 6: sp.mammal:*
- 7: sp.mhc:*
- 8: sp.organelle:*
- 9: sp.phase:*
- 10: sp.plant:*
- 11: sp.rodent:*
- 12: sp.virus:*
- 13: sp.vertebrate:*
- 14: sp.unclassified:*
- 15: sp.virus:*
- 16: sp.bacteriap:*
- 17: sp.bacteriap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1113	77.3	207	4	Q9BX74 Homo sapiens
2	941.5	65.4	208	11	Q99MB3 Q99MB3 mus musculus
3	338	23.5	178	5	Q9W2H1 Q9W2H1 drosophila
4	278	19.3	329	5	Q95P78 Q95P78 caenorhabditis
5	200	13.9	221	4	Q9H651 Q9H651 homo sapien
6	200	13.9	247	4	Q9BN9 Q9BN9 homo sapien
7	198.5	13.8	261	11	Q8BJ83 Q8BJ83 mus musculus
8	196	13.6	230	11	Q9D156 Q9D156 mus musculus
9	182	12.6	284	5	Q9U4H5 Q9U4H5 drosophila
10	172	12.0	80	4	Q9H046 Q9H046 homo sapien
11	169.5	11.8	195	5	Q9S025 Q9S025 caenorhabditis
12	136.5	9.5	214	4	Q9BX73 Q9BX73 homo sapien
13	125	8.7	172	5	Q9V786 Q9V786 drosophila
14	121	8.4	149	4	Q9BSR6 Q9BSR6 homo sapien
15	121	8.4	171	4	Q8N0X9 Q8N0X9 homo sapien
16	117	8.1	106	2	Q9S022 Q9S022 borrelia bu

17	115.5	8.1	213	11	Q9CWL9 Q9CWL9 mus musculus
18	116.5	8.1	213	11	Q8R014 Q8R014 mus musculus
19	116.5	8.1	213	11	Q8BJJ1 Q8BJJ1 mus musculus
20	107	7.4	106	16	Q8F808 Q8F808 leptospira
21	105.5	7.3	142	16	Q8DJB9 Q8DJB9 synchococcus
22	105	7.3	89	2	Q9S015 Q9S015 borrelia bu
23	104	7.2	159	16	Q8F099 Q8F099 corynebacter
24	98	6.8	1092	5	Q964R2 Q964R2 theileria t
25	97	6.7	137	16	Q912P1 Q912P1 staphylococ
26	95.5	6.6	913	4	Q7S061 Q7S061 homo sapien
27	95	6.6	309	16	Q9RS03 Q9RS03 deinococcus
28	92.5	6.4	76	16	Q8EVR7 Q8EVR7 mycoplasma
29	92	6.4	456	12	Q66668 Q66668 equine herp
30	92	6.4	720	5	Q9Y5W4 Q9Y5W4 drosophila
31	92	6.4	720	5	Q961T2 Q961T2 drosophila
32	91	6.3	1324	12	Q9J3E7 Q9J3E7 murine hepa
33	90.5	6.3	151	11	Q8K181 Q8K181 mus musculus
34	90	6.3	204	10	Q8R2E7 Q8R2E7 otyza sativ
35	89.5	6.2	690	10	Q22512 Q22512 glycine max
36	89	6.2	330	12	Q98339 Q98339 murine hepa
37	88	6.1	1324	12	Q39227 Q39227 murine hepa
38	87.5	6.1	400	16	Q9LIE7 Q9LIE7 streptomyce
39	87.5	6.1	547	16	Q9RZJ4 Q9RZJ4 deinococcus
40	87.5	6.1	1207	5	Q21535 Q21535 caenorhabdi
41	87	6.0	1274	10	Q9M680 Q9M680 dunaliella
42	87	6.0	1321	12	Q9J3F1 Q9J3F1 murine hepa
43	86.5	6.0	364	5	Q9VXJ5 Q9VXJ5 drosophila
44	86.5	6.0	377	16	Q9KYT3 Q9KYT3 streptomyce
45	86.5	6.0	381	17	Q978F6 Q978F6 thermoplasm

ALIGNMENTS

RESULT 1

Q9BX74 Q9BX74 PRELIMINARY; PRT; 207 AA.
AC Q9BX74:
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Beta-amyloid binding protein.
GN BAP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=21276355; PubMed=11278849;
RX KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
Containing a G Protein Activation Module";
RI J. Biol. Chem. 276:18748-18756(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21276355; PubMed=11278849;
RA KAJKOWSKI E.M., LO C.F., NING X., WALKER S., SOFIA H.J., WANG W.,
RA EDLIS W., CHANDA P., WAGNER E., VILLE S., RYAN K., McHERDREY-RINDE B.,
RA SMITH S.C., WOOD A., RHODES K.J., KENNEDY J.D., BARD J.,
RA JACOBSEN J.S., OZENBERGER B.A.;
RT "beta-Amyloid

```
Db 1 MAAAMPSPGSAPEAVTARLVGLWVSVTTGPGAVATISAGCESLKCED.KVGQYICKD 60
123 PKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHTFTGNEVGFKPK.SCRNVNGYS 182
61 PKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHTFTGNEVGFKPK.SCRNVNGYS 120
183 YKVAVALSLFLGWLGCADRFYLGYPALGLLKFTCTVGCIGISLIDFILISMQIVGPGSDGS 242
121 YKVAVALSLFLGWLGCADRFYLGYPALGLLKFTCTVGCIGISLIDFILISMQIVGPGSDGS 180
243 YIIDYGGTGLRLSLTNETFRKTOLYP 269
181 YIIDYGGTGLRLSLTNETFRKTOLYP 207

RESULT 2
Q99MB3 PRELIMINARY: PRT: 208 AA.
AC Q99MB3;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Beta-amyloid binding protein.
GN BPP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=21276355; PubMed=11278849;
RA Kajkowski E.M., Lo C.F., Ning X., Walker S., Sofia H.J., Wang W.,
RA Edris W., Chanda P., Wagner E., Ville S., Ryan K., McHenry-Rinde H.,
RA Smith S.C., Wood A., Rhodes K.J., Kennedy J.D., Bard J.,
RA Jacobsen J.S., Ozenberger B.A.;
RT "beta-Amyloid peptide-induced Apoptosis Regulated by a Novel Protein
RT Containing a G Protein Activation Module."
RL J. Biol. Chem. 276:18748-18756(2001).
DR EMBL; AF353993; AAK35067.1; -.
DR MGD; MG1:2137022; Bbp.
SQ SEQUENCE 208 AA: 22271 MW: 9147932163F4F04C CRC64;

Query Match 65.4%; Score 941.5; DB 11; Length 208;
Best Local Similarity 85.1%; Pred. No. 8.4e-79;
Matches 177; Conservative 10; Mismatches 20; Indels 1; Gaps 1;

QY 63 MAAAMPSPGSAPEAVTARLVGLWVSVTTGPGAVATISA-OGESLKCEDLKVGQYICK 12;
Db 1 MAAAMPGRASPAAGPPGLRLTLVTVAAHGCAAGVGGETPKCEDLRVGQYICK 60
122 DKINDATQEPVNCNTYTAHVSCFPAPNITCKDSSGNETHTFTGNEVGFKPK.SCRNVNGY 181
61 EPKINDATQEPVNCNTYTAHVSCFPAPNITCKDLSGNETHTFTGSEVGLFKPKISCRNVNGY 120
182 SYKVAVALSLFLGWLGCADRFYLGYPALGLLKFTCTVGCIGISLIDFILISMQIVGPGSDGS 241
121 SYKVAVALSLFLGWLGCADRFYLGYPALGLLKFTCTVGCIGISLIDFILISMQIVGPGSDGS 180
242 SYIIDYGGTGLRLSLTNETFRKTOLYP 269
181 SYIIDYGGTGLRLSLTNETFRKTOLYP 208

RESULT 3
Q9W2H1 PRELIMINARY: PRT: 178 AA.
AC Q9W2H1;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE CG10795 protein (LD27358P).
```

```
CG10795.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew K.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hootin D., Houston K.A., Howland T.J., Wei M.-H., Ibeqam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mallet B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzzy D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclele J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M.D., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao C., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
Science 287:2185-2195(2000).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuno J., Paclele J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003453; AAF46720.1; -.
DR EMBL; AY061343; AAL28891.1; -.
DR FlyBase; FBgn0034626; CG10795.
SQ SEQUENCE 178 AA: 19896 MW: 17C41166607ACC03 CRC64;

Query Match 23.5%; Score 338; DB 5; Length 178;
Best Local Similarity 42.6%; Pred. No. 3e-23;
Matches 69; Conservative 30; Mismatches 49; Indels 14; Gaps 5;
```

```
QY 107 SLKCEDLK-VGQYICKDP---KINDATQEPVNCNTY-TAHVSCFPAPNITCKDSSGNETH 161
:: | :: :||::| :: | :: | :: | :: | :: | :: | :: | :: | :: | :: |
DB 20 NVDCNELQMGGFCPPDPARGOIDPKTQOLAGCTREGRCYWCIAANEINCTF-TGNAT- 77
QY 162 FTGNEVGFKPKISCRNVNGYSYKVAVALSLFLGWLGCADRFYLGYPALGLLKFTCTVGC 221
:: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: |
DB 78 -----FTRVPCKWYNGYHLDITLLLSVFLGWFVGVDRLFYLGPGGLLAKFTLGGMFL 130
```

```

RP SEQUENCE FROM N.A.
RA Kajkowski E.M., Lo C.P., Ning X., Walker S., Sofia H.J., Wang W.,
RA Edris S.W., Chanda P., Wagner E., Vile S., Ryan K., McHendry-Rinde B.,
RA Smith S.C., Wood A., Rhodes K.J., Kennedy J.D., Bard J.,
RA Jacobsen J.S., Ozenberger B.A.;
RA "Beta-amyloid peptide-induced apoptosis regulated by a novel protein
RT containing a G protein activation module.";
RI J. Biol. Chem. 0:0-0(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RM EMBL: AK026257; BAB15415.1; -.
RM EMBL: AF353992; AAK3066.1; -.
RM EMBL: BC008873; AAH08873.1; -.
RW Hypothetical protein
SQ SEQUENCE 221 AA; 24410 MW; 92151D6EF636D74 CRC64;

Query Match 13.9%; Score 200; DB 4; Length 221;
Best Local Similarity 45.7%; Pred. No. 2.1e-10;
Matches 53; Conservative 12; Mismatches 39; Indels 12; Gaps 5;

QY 135 CTNYTA--HVSC-----FPAPNITCKDSSGNETHFTGNEVGFFKPISCRNVNGYSYKVAVA 188
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
112 CTNSTSCMTVSCPQRYPAP-NCTVRD-----HVHCLGNRT-FPKMLYCNTTGGYKWKSTALA 165

QY 189 LSLFLGWLGAADRYLGYALGALLKFKCTVSGCGIGSLIDFILISMQIVGSPDGSYYI 244
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
166 LSLITLGGFGADRYLQWREGLGKLFSGGLGIWTLIDVLLIGVGVGPDGSLYI 221

RESULT 6
Q5BRN9 PRELIMINARY; PRT; 247 AA.
AC Q5BRN9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Similar to hypothetical protein FLJ22604.
CS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CX Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CQ NCBI_TaxID=9606;
CN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RM EMBL: BC006150; AAH06150.1; -.
RW Hypothetical protein
SQ SEQUENCE 247 AA; 27161 MW; CE1DGD9C53DDF73C CRC64;

Query Match 13.9%; Score 200; DB 4; Length 247;
Best Local Similarity 45.7%; Pred. No. 2.4e-10;
Matches 53; Conservative 12; Mismatches 39; Indels 12; Gaps 5;

QY 135 CTNYTA--HVSC-----FPAPNITCKDSSGNETHFTGNEVGFFKPISCRNVNGYSYKVAVA 188
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
138 CTNSTSCMTVSCPQRYPAP-NCTVRD-----HVHCLGNRT-FPKMLYCNTTGGYKWKSTALA 191

QY 189 LSLFLGWLGAADRYLGYALGALLKFKCTVSGCGIGSLIDFILISMQIVGSPDGSYYI 244
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
192 LSLITLGGFGADRYLQWREGLGKLFSGGLGIWTLIDVLLIGVGVGPDGSLYI 247

RESULT 7
Q8B383 PRELIMINARY; PRT; 261 AA.
AC Q8B383;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

```



```

DB 90 PATNCSTSTKLLTKCSAHSVICHQORNEYKRIPCNCWSSCSWIKTILMSVVLGGGFA 149
QY 199 DREYLGYPALCGLLAKCTVCGFCIGSLIDFILISMOIVCPGSGSSVI 244
DB 150 DREYLGWLKSAIGLKFSGGGLVMTLVDVLIANGYIKPYDGSMTY 195

RESULT 12
Q9BX73 PRELIMINARY; PRT; 214 AA.
AC Q9BX73;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE BBP-like protein 1.
GN BLP1.
OS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=21276355; PubMed=11278849;
RA Rajkowsky E.M., Lo C.F., Ning X., Walker S., Sofia H.J., Wang W.,
RA Edris W., Chanda P., Wagner E., Ville S., Ryan K., McHenry-Rinde R.,
RA Smith S.C., Wood A., Rhodes K.J., Kennedy J.D., Bard J.,
RA Jacobsen J.S., Ozenberger B.A.;
RT "beta-Amyloid Peptide-induced Apoptosis Regulated by a Novel Protein
RT Containing a G Protein Activation Module.";
RL J. Biol. Chem. 276:19748-19756(2001).
DR EMBL; AF353991; AAK35065.1; -.
SQ SEQUENCE 214 AA; 22871 MW; BB928712AF2F78AB CMC64.

Query Match 9.5%; Score 136.5; DB 4; Length 214;
Best Local Similarity 27.8%; Pred. No. 0.00014;
Matches 58; Conservative 21; Mismatches 83; Indels 47; Gaps 10;

QY 59 SVSKMAAAMP-----SGPSAPEAVTARLVGVLMFVSVTIGPGAVATSAGGES--LKCED 112
DB 33 SHSQNATAPELTISAGAOPE-----GPGASASWEYGDHPSPVLASY 75
QY 113 LKVGQYICKDP--KINDAT--QE-PVNCNTNTAH-----VSCFPAPNITCKDSSGN 156
DB 76 LPDEFIECDVDHVGNGNATASQELQYGLKFGGAYSDVENTSVQCHALDIEC---ASP 132
QY 159 ETHFTGNVEGFKPISCRNNGYSKVAVALSLFLGWLGDREVLYGYPALGLLFCVIGF 216
DB 133 RTFLREN-----KP--CKYTGHTFHTLLYSFFLGGCGVDRFCLGHTGTAVGKILLIGG 185
QY 219 CGIGSLIDFILISMOIVGPGSGSSVIIDY 247
DB 186 LG1WMFVLDLILLTGLGLMPDGSNNCTIVY 214

RESULT 13
Q9VY86 PRELIMINARY; PRT; 172 AA.
AC Q9VY86;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE CG11103 protein (LP03404p).
GN CG11103.
OS Drosophila melanogaster (Fruit fly).
CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Eteryceta;
CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OX Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne C.D.,

```

```

Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Ahril J.F., Agbayani A., An H.-J., Andrews-Frannkoch C., Baldwin D.,
Bailey R.M., Basu A., Baxendale J., Bayraktaroglu L., Basley E.M.,
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Dowies M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
Fosier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison C.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mount S.M., Moy M., Murphy B., Murphy L., Muzzy D.M., Nelson D.L.,
Nelson D.R., Nelson X.A., Nixon K., Nusskern D.R., Paclab J.M.,
Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen R.,
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Svirskas R., Tector C., Turner R., Venter F., Wang A.H., Wang X.,
Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong E.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.G.,
Gibbs S.A., Myers E.W., Rubin G.W., Venter J.C.;
X. "The genome sequence of Drosophila melanogaster.";
Science 287:2185-2195(2000).
[2]
SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busan D.A.,
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
Dodson K., Dorset V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
Paclab J., Paragay V., Park S., Patel S., Pfeiffer B.,
Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
Williams S.M., Zaveri J.S., Smith H.G., Venter J.C., Rubin G.M.;
"Sequencing of Drosophila melanogaster genome.";
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
"Annotation of Drosophila melanogaster genome.";
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[5]
SEQUENCE FROM N.A.
RA Flybase;
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

```

